Atomic Energy Control Board
P.O. Box 1046
Ottawa, Canada
K1P 5S9

Commission de contrôle de l'énergie atomique
C.P. 1046
Ottawa, Canada
K1P 5S9

INFO-0241

MODELLING OPERATOR COGNITIVE INTERACTIONS IN NUCLEAR POWER PLANT SAFETY EVALUATION

by

J.W. Senders, N. Moray,
A. Smiley and A. Sellen
MMS Man-Machine Systems
Consultants, Inc.

A research report prepared for the
Atomic Energy Control Board
Ottawa, Canada

Project No. 84.1.21

August 28, 1985

Canada

Research report

MODELLING OPERATOR COGNITIVE INTERACTIONS IN NUCLEAR POWER PLANT SAFETY EVALUATION

ABSTRACT

This report presents the results of a research study conducted in support of the human factors engineering program of the Atomic Energy Control Board of Canada. The overall objectives of the study were to review methods which are applicable to the analysis of control room operator cognitive interactions in nuclear plant safety evaluations and to indicate where future research effort in this area should be directed.

An understanding of the nature and origins of operator errors is of the greatest importance, because in well-designed modern industrial systems most system faults and failures are due in some way to human error, not to mere hardware failure.

This report is based on an exhaustive search and review of the literature on NPP operator error, human error, human cognitive function, and on human performance. A number of methods which have been proposed for the estimation of data for probabilistic risk analysis have been examined and have been found wanting. None addresses the problem of diagnosis error per se. Virtually all are concerned with the more easily detected and identified errors of action. None addresses underlying cause and mechanism. It is these mechanisms which must be understood if diagnosis errors and other cognitive errors are to be controlled and predicted.

The literature does offer a long list of task variables which influence cognitive performance. Although the information does not allow complete quantification, it does give guidance for cognitive error rate reduction.

No definition of error is completely satisfactory and widely accepted. This has led to difficulty in creating acceptable taxonomies with which to classify errors. Errors of diagnosis are undistinguishable from any other kinds of error unless one resorts to careful debriefing of operators, or attributes an overt error to the diagnostic process by modelling error mechanisms and eliminating alternative explanations. The models to be found in the literature do not allow prediction of error nor do they provide a basis for a systematic investigation of error mechanisms. We present a number of taxonomies and models and examine their shortcomings.

We have attempted to overcome the deficiencies of earlier work and have constructed a model/taxonomy, EXHUME, which we consider to be exhaustive. This construct has proved to be fruitful in organizing our thinking about the kinds of error that can occur and the nature of self-correcting mechanisms, and has guided our thinking in suggesting a research program which can provide the data needed for quantification of cognitive error rates and of the effects of mitigating efforts. In addition a preliminary outline of EMBED. a causal model of error, is given based on general behavioural research into perception, attention, memory, and decision making.

Finally, we propose a research plan and define some of the necessary characteristics of a facility that would be required.

RÉSUMÉ

Le présent rapport donne les résultats d'une recherche effectuée à l'appui du programme d'ergonomie de la Commission de contrôle de l'énergie atomique du Canada. L'étude visait principalement à examiner des méthodes qui s'appliquent à l'analyse des interactions d'ordre cognitif entre les opérateurs de salle de commande et les installations dans le cadre de l'évaluation de la sûreté des centrales nucléaires et à indiquer dans quelle direction pourraient être orientées les futures recherches dans ce domaine.

Il est très important de comprendre la nature et l'origine des erreurs d'opérateurs, car, dans des systèmes industriels modernes bien conçus, la plupart des incidents sont attribuables non pas seulement à une défaillance du matériel, mais à l'erreur humaine sous une forme quelconque.

Le rapport est fondé sur une recherche et un examen exhaustifs de la documentation sur les erreurs d'opérateurs dans les centrales nucléaires, l'erreur humaine, la fonction cognitive et le rendement de l'être humain. Plusieurs méthodes proposées pour l'évaluation des données aux fins de l'analyse probabiliste du risque ont été examinées et jugées insuffisantes. Aucune n'aborde le problème de l'erreur du diagnostic comme tel. Presque toutes s'intéressent aux erreurs d'exécution qui sont plus faciles à déceler et à identifier. Aucune ne porte sur les causes et les mécanismes sous-jacents. Ce sont ces mécanismes qui doivent être compris si l'on veut imiter et prévoir les erreurs de diagnostic et les autres erreurs d'ordre cognitif.

La documentation offre une liste de variables des tâches qui influencent le rendement cognitif. Bien que cette information ne permette pas de faire une quantification complète, elle donne des conseils pour réduire l'erreur d'ordre cognitif.

Aucune définition de l'erreur n'est entièrement satisfaisante et largement acceptée, d'où la difficulté de créer des taxonomies acceptables qui permettent de classer les erreurs. Il est difficile de distinguer les erreurs de diagnostic des autres types d'erreurs, à moins d'avoir recours à des comptes rendus précis des opérateurs ou d'attribuer une erreur observée au processus de diagnostic en modelant les mécanismes de l'erreur et en supprimant les autres explications. Les modèles notés dans la documentation ne permettent pas de prévoir l'erreur, pas plus qu'ils ne fournissent de base pour une étude systématique des mécanismes de l'erreur. Nous présentons plusieurs taxonomies et modèles, et examinons leurs lacunes.

Nous avons tenté de corriger les insuffisances des travaux antérieurs et avons élaboré un modèle ou une taxonomie intitulé EXHUME, que nous considérons comme exhaustif. Cette tâche nous a permis d'organiser nos idées à propos des types d'erreurs qui peuvent survenir et sur la nature des mécanismes d'autocorrection, et elle nous a amenés à suggérer un programme de recherche qui pourrait fournir les données nécessaires pour quantifier le taux des erreurs d'ordre cognitif et évaluer les effets des actions visant à les atténuer. De plus, nous donnons un aperçu préliminaire du modèle causal de l'erreur, EMBED, qui est fondé sur des recherches générales sur la perception, l'attention, la mémoire et la prise de décision.

Finalement, nous proposons un plan de recherche et définissons quelques-unes des caractéristiques de l'installation requise pour effectuer une telle recherche.

Table of Contents

EXECUTIVE SUMMARY Page 6

1.0 INTRODUCTION Page 7

2.0 DEFINITION OF COGNITIVE INTERACTIONS Page 9

2.1 Cognitive Behaviour 9
2.2 Mental Models 10
2.3 Examples of Cognitive Behaviour in Nuclear Power Plants 12
2.4 An Outline of Cognitive Mechanisms 13

3.0 LITERATURE REVIEW Page 17

3.1 The Scope of the Review 17
3.2 The Nature of the Human as an Information Processor and Decision Maker 17
3.2.1 Dynamic Memory 17
3.2.2 Decision Making Ability 19
3.2.3 Response to Faults and Emergencies 21
3.2.4 Fault Detection and Diagnosis and Human Error 22
3.3 Mental Models and the Prediction of Error 23
3.3.1 Characteristics of Mental Models 23
3.3.2 Methods for Analysis of Mental Models 24
3.4 Concepts and Models 26
3.4.1 The Nature of Error 26
3.4.2 Concepts of Error 27
3.4.3 Conceptual Models 29
3.4.4 Quantitative Models 33
3.4.5 Overview of Models 40
3.5 Taxonomies 40
3.5.1 The Nature of Taxonomies 40
3.5.2 Behaviour Oriented Taxonomies 41
3.5.3 Task Oriented Taxonomies 49
3.6 Cognitive Engineering 50
3.7 The Assessment of Behaviour Involving Cognitive Error 53
3.7.1 Methods of Assessment 53
3.7.2 Significant Event Reports 53
3.7.3 Nuclear Power Plant Simulation Studies 53
3.7.4 Simulations of Plants Other Than NPPs 54
3.7.5 Laboratory Studies 55
3.7.6 Verbal Reports 57

4.0 PROPERTIES OF HUMAN INFORMATION PROCESSING WHICH TEND TO PRODUCE ERRORS IN COGNITIVE ACTIVITIES Page 59

4.1 Human Information Processing and System Design 59
4.2 Memory 59
4.3 Attention 59
4.4 SpeedAccuracy Tradeoff 59
4.5 Cognitive Tunnel Vision 60
4.6 Statistical Inference 60
4.7 Logical Inference 60
4.8 Decision Making 61

5.0 SYNTHESIS Page 62

5.1 The Definition of Error 62
5.2 Error Taxonomies 62
5.3 The Question of Whether Human Error is Caused 63
5.4 On the Modelling of Human Error 63
5.4.1 Levels of Analysis 63
5.4.2 Exhaustive Model of Human Error (EXHUME) 65
5.5 The Problem of Detecting Cognitive Error 68
5.6 MultiStage Closed Loop Model of Behaviour 69
5.7 The Information Required for Quantitative Prediction 70
5.8 Error Model Based on Empirical Data (EMBED) 70

6.0 RESEARCH PLAN Page 75

6.1 Alternative Ways of Gathering Data 75
6.1.1 L.E.R.'s and S.E.R.'s 75
6.1.2 Expert Opinion 75
6.2 Research Plan for Error Data Collection 76
6.3 Research Plan for EMBED 77
6.4 Research Plan for Quantification of Cognitive Error Mechanisms. 78

7.0 CONCLUSIONS Page 79

8.0 ANNOTATED BIBLIOGRAPHY Page 80

9.0 REFERENCES Page 95

10.0 LIST OF ABBREVIATIONS Page 107

11.0 ACKNOWLEDGEMENTS Page 109

List of Figures

Figure 1. A Flow Chart of Human Information Processing. Page 13

Figure 2. Process Control Skills of the Operator. (From Rasmussen, 1983) Page 14

Figure 3. Flow Diagram Showing the Basic Components of the Composite-Skill Model. (From Reason, 1977) Page 29

Figure 4. Algorithm for Classifying Actions Not-As-Planned. (From Reason, 1977) Page 31

Figure 5. A Model of Error Production. Page 32

Figure 6. A Model of Human Decision Sequence. (From Rasmussen, 1981) Page 34

Figure 7. Example of the Operator Action Tree. Page 37

Figure 8. Multifacet Taxonomy For Description and Analysis of Events Involving Human Malfunction. (From Rasmussen, 1981) Page 42

Figure 9. Senders' Classification Scheme For Error. Page 43

Figure 10. The Berliner Taxonomy. (From Berliner et. al, 1964) Page 44

Figure 11. Altman/Moray Scheme for Classification of Error. (From Moray, 1981) Page 47

Figure 12. EXHUME: A MultiStaged Closed Loop Model of Behaviour. Page 66

Figure 13. EMBED: Error Model Based on Empirical Data. Page 73

List of Tables

Table 1. Examples and Relative Frequencies of Actions Not-As-Planned. (From Reason, 1977) Page 45

Table 2. Categorization of Action Slips. (From Norman, 1980) Page 48

Table 3. Taxonomy of NPP Operator Error. (From Beattie and Iwasa-Madge, 1984) Page 51

Table 4. Analysis and Classification of Human Errors in Troubleshooting Live Aircraft Power Plants. (From Johnson and Rouse, 1982) Page 56

Table 5. Information From Verbal Reports. (From Bainbridge, 1979) Page 58

EXECUTIVE SUMMARY

Modelling Operator Cognitive Interactions In Nuclear Power Plant Safety Evaluation.

This report presents the results of a research study conducted in support of the human factors engineering program of the Atomic Energy Control Board of Canada. The overall objectives of the study were to review methods which are applicable to the analysis of control room operator cognitive interactions in nuclear plant safety evaluation and to indicate where future research effort in this area should be directed.

An understanding of the nature and origins of operator errors is of the greatest importance, because in well designed modern industrial systems most system faults and failures are due in some way to human error, not to mere hardware failure.

This report is based on an exhaustive search and review of the literature on Nuclear Power Plant (NPP) operator error, human error, human cognitive function, and on human performance. A number of methods which have been proposed for the estimation of data for probabilistic risk analysis have been examined and have been found wanting. None addresses the problem of diagnosis error per se. Virtually all are concerned with the more easily detected and identified errors of action. None addresses underlying cause and mechanism. It is these mechanisms which must be understood if diagnosis errors and other cognitive errors are to be controlled and predicted. The literature does offer a long list of task variables which influence cognitive performance. Although the information does not allow complete quantification, it does give guidance for cognitive error rate reduction.

No definition of error is completely satisfactory and widely accepted. This has led to difficulty in creating acceptable taxonomies with which to classify errors. Errors of diagnosis are indistinguishable from any other kinds of error unless one resorts to careful debriefing of operators, or attributes an overt error to the diagnostic process by modelling error mechanisms and eliminating alternative explanations. The models to be found in the literature do not allow prediction of error nor do they provide a basis for a systematic investigation of error mechanisms. We present a number of taxonomies and models and examine their shortcomings.

We have attempted to overcome the deficiencies of earlier work and have constructed a model/taxonomy, EXHUME, which we consider to be exhaustive. This construct has proved to be fruitful in organizing our thinking about the kinds of error that can occur and the nature of self-correcting mechanisms, and has guided our thinking in suggesting a research program which can provide the data needed for quantification of cognitive error rates and of the effects of mitigating efforts. In addition a preliminary outline of EMBED, a causal model of error, is given based on general behavioural research into perception, attention, memory, and decision making. Finally, we propose a research plan and define some of the necessary characteristics of a facility that would be required.

DISCLAIMER

The Atomic Energy Control Board is not responsible for the accuracy of the statements made or opinions expressed in this publication and neither the Board nor the authors assume liability with respect to any damage or loss incurred as a result of the use made of the information contained in this publication.

1.0 INTRODUCTION

This report was prepared for the Atomic Energy Control Board during the period June 1984 to June 1985. The purpose of the report is to assist the AECB in understanding the origins and nature of operator error, with particular emphasis on cognitive errors; to indicate ways in which cognitive error rates may be classified, modelled, and predicted; and to indicate areas where research is both desirable and feasible. Cognitive errors are errors of planning, judgement, diagnosis, etc. which occur within the brain of the operator and which may have no observable behavioural concommitants.

An understanding of the nature and origins of operator errors is of the greatest importance, because in well designed modern industrial systems most system faults and failures are due in some way to human error, not to mere hardware failure.

Relevant literature has been reviewed, and interviews carried out both with individuals who had considerable experience on control room operators and with experts on human error research. Research on the psychology of human error in general has been examined as well as work more specifically related to the nuclear industry. In addition general literature on the nature of human information processing has been cited where it is relevant to an understanding of error, its nature and origins.

The report reviews taxonomies and models of human error, and methods which have been proposed for predicting error rates and the occurrence of specific errors. Particular attention is paid to the distinction between errors of behaviour and cognitive errors.

Three kinds of models are suggested. The first is one based on the statistical analysis of observed errors. Such a model is not feasible at present because of an absence of data. The second is a model derived from the inherent logical structure of information processing. The third is derived from empirically determined facts about factors affecting the efficiency of human information processing.

The importance of cognitive errors will increase with the development of advanced control rooms. The application of supervisory control decreases the behavioural interaction between operator and system. Already, for example, CANDU NPP operators perform far fewer manual control actions than do U.S. PWR operators. The opportunities for behavioural errors are correspondingly reduced. On the other hand, the opportunities for cognitive error increase. Major tasks of the operator include monitoring system status, making decisions, and keeping a "mental model" of the plant status up to date. In the absence of any feedback (due to the lack of interaction with the system), the occurence of cognitive error is more likely to go unnoticed. Consequently, if the operator is suddenly required to intervene manually in response to an abnormal plant state the probability of erroneous action must increase. The importance of understanding the cognitive origins of error is that an operator may take action which is correct in the light of an internal (mental) cognitive judgement. Even though the behaviour is incorrect from a plan performance point of view, it is entirely correct given the initial cognitive judgement. The behaviour is not in error. The cognitive judgement is.

To support the decision-making and diagnostic behaviour of an operator requires a control room interface which matches the information processing abilities of an operator. Similarly procedures must be written and training carried out in an appropriate way. Conventional human factors and ergonomics have concentrated on understanding how to minimize behavioural error. The present report is a first step to performing the same service with respect to the unobservable mental actions which are increasingly critical and prevalent in human-machine systems such as NPPs.

2.0 DEFINITION OF COGNITIVE INTERACTIONS

2.1 Cognitive Behaviour

Many definitions of "cognitive behaviour" were found in the literature. Wreathall (1983) uses the term for behaviour that involves structuring information, conceptualizing cause-effect relationships, and responding to a situation as cognitive behaviour. Norman (1981) prefers to speak of intelligent behaviour, and he defines cognition as the study of all mental phenomena between the sensory and the motor aspects of behaviour. It is clear that the domain of mental activity that is important for human reliability in this context is beyond the realm of simple stimulus-response processes. That is, it is no longer fruitful to regard the human as a black box and concentrate only on observable inputs and outputs. Rather, we know that humans engage in complex, internal mental activities such as planning, diagnosing, and interpreting the environment. Behaviour is often "top-down", or goal oriented and purposive. The term cognitive interaction stresses the dynamic nature of the human - machine interaction, as well as the fact that not only does the environment act upon the human, but the human is an active modifier of the environment. Cognitive behaviour is behaviour which analyses, judges, and transforms information received from the environment; behaviour which makes use of memory to interpret such information; which discovers meaning or imposes meaning; which predicts, interpolates and extrapolates; which decides and chooses goals, plans and courses of acton. Cognitive behaviour, to use Bruner's phrase, goes "beyond the information given" (Bruner, 1973).

In modelling the cognitive behaviour of operators interacting with a complex system, many researchers (psychologists in particular) choose to view the human as an information processor and accordingly, break down behaviour into stages of processing. (See for example, the proceedings of a Workshop on Cognitive Modelling, NUREG 3114, 1982.) Two important concepts need to be considered as central to an understanding of operator cognitive behaviour.

1) Cognitive behaviour is a function of operator knowledge. Knowledge is structured, often incomplete, and often erroneous.

2) People use different levels of knowledge and exhibit a large amount of flexibility in what level of knowledge they bring to bear on problems depending on what their goals are.

2.2 Mental Models

Most researchers believe in and invoke mental models formed in the "actor" and frequently updated, to explain cognitive interaction. The "mental model" is a structured unit or combination of units of knowledge which govern how we diagnose, interpret, and select actions for dealing with external events, as well as enabling us to predict future events. The idea that knowledge is both organized and exists as levels or packages of information is the basis for the psychological literature on schemata, mental models, frames, and scripts. For a good discussion of these theoretical concepts, see NUREG 3114 (1982). All of these terms refer to the idea that knowledge provides a context within which meaning is attached to stimuli in the environment (Braune and Forshay, 1983). We can think of a mental model as knowledge which is called into memory. The character of a mental model is such that is has certain "default" values. That is, there are concepts and beliefs represented in the model which can account for information that is already known about the environment. Thus, where uncertainty exists, the operator will seek out the unknown information. This also means that hypothesis formation, planning behaviour, and diagnosis will be based on these implicit assumptions or default values when explicit values cannot be obtained.

A model is always something less than the thing modelled. The operator may be overloaded, and so try to diagnose the plant state using fewer variables then are really required. The operator may use all the relevant variables, but may be unable to understand the currently displayed concatenation of information and so make use of other values which in the past he succeeded in understanding. In these cases default values are not used; instead the true state of the plant is "modelled" more simply by the operator.

With respect to nuclear power plant operators, Sheridan (1982) lists several kinds of knowledge that must be incorporated into the operator's mental model. They are:

1. hardware configuration of the plant and control panel

2. dynamic behaviour of the plant variables

3. nominal operating procedures (rules of thumb included)

4. control criteria and objectives

5. abstract event relations

6. administrative procedures and management policies

With respect to levels of knowledge that the operator may bring to bear to cope with task demands, Rasmussen (1982), characterizes them on a continuum of abstraction ranging from the representation of the physical form of the plant to the more abstract and holistic level of the functional purpose of goals and intentions. The operator may use any knowledge level that is deemed most appropriate to act upon and assess any situation. For example, when looking for information on a control panel, the operator may invoke a schema for the layout of the control room and then bring back into memory the schema containing high level goals to provide the purpose of seeking out that information. A more detailed taxonomy for categorizing mental models can be found in Rasmussen (1979).

2.3 Examples of Cognitive Behaviour in Nuclear Power Plants

There are two sources of examples of cognitive behaviour in NPPs. One is operating procedures, particularly emergency operating procedures (EOPs); the other is the behaviour of operators and their reports about their behaviour. Although there are always problems involved in analyzing verbal reports either during operations (protocol analysis) or afterwards (interviews) there is no doubt that such methods provide important and valid insights (Bainbridge, 1974, 1981a,b; Ericksen and Simon, 1984). "Verbal reports" and "verbal behaviour" in this document imply that words, whether spoken or written, are used by the operator to describe what he is doing or has done. The word "verbal" is used by contrast with behaviour such as the operation of switches, push-buttons, etc. It implies that the behaviour has meaning which cannot be discovered except by examining what is said or written.

It is usually the intention of EOPs to reduce cognitive behaviour to a minimum, because of its uncertain outcome and because of our lack of knowledge about how to support cognitive behaviour. EOPs should ideally direct the operator to examine a particular display and observe the value. Depending on that value some action may have to be taken. The action may be to affect the plant state by operating a switch, opening a valve, etc., or it may be to take another reading from the same or another instrument. If the EOP consists of statements of the kind:

"If the temperature shown on gauge SM159 is greater than 170 degrees, then turn on pump XJ74."

"If valve MP110 is closed then turn off pump 14PP7B."

"If the rpm shown on meter RM6 is greater than 600 and less than 650 then change the setting on the valve control V76B to 7",

then cognitive behaviour is minimized, since information processing is reduced to perception and motor response. (Even here, however, as will be shown, there are certain possible cognitive interactions.)

However, if EOPs contain instructions of the following kind:

"If the temperature is rising rapidly ... open valve V2R7..."

"If it is estimated that the pump can be repaired in 20 minutes ... restart the computer..."

"If the water level is too low ... activate the pump."

"If the difference between the hot and cold leg temperatures is excessive...",

then cognitive behaviour is usually involved. This is because judgement must be applied to data; working memory must be used to compare data and predictions made from present to future plant state. One aim of good EOPs and control room design is to minimize cognitive behaviour.

Discussions with operators reveal that they are aware of the extent to which cognitive processes rather than explicit EOPs control their behaviour. When a transient occurs a large amount of evidence about plant state appears on CRTs, annunciators, etc. In order to decide which EOP to follow, some interpretation is necessary. It is often reported that "knowledge-based behaviour" must be used to supplement the rule-based or skill-based behaviour supported and expected by EOPs. Operators see their behaviour as determined both by explicit procedures and also by their understanding of how the system works. This is of course particularly true for abnormal conditions for which there are not EOPs. Furthermore, operators commonly speak of "seeing" the physical plant through or behind the board, suggesting that they operate with a mental model, and such a model, being an abstraction or mental construction, is certainly cognitive. Indeed operators sometimes complain that one problem with so called high fidelity simulators is that they are sufficiently different from the real plant that the mental model is not effective.

An obvious example of cognitive behaviour is the response of operators to a false alarm. An illuminated annunciator may indicate a genuine plant fault or may indicate a faulty annunciator or faulty sensor. Once the operator has decided that the latter is the case, his attitude, judgements, decision, and behaviour toward the system change. The decision not to carry out corrective action because the annunciation is faulty is obviously a cognitive act.

2.4 An Outline of Cognitive Mechanisms

At a later stage (Section 5) of this report a flow chart model of operator behaviour will be presented which will relate different kinds of error to different stages of information processing. It will be an information processing model of error. Logically prior to such a model is a model of information processing itself, (see Figure 1) which will allow the reader to relate psychological studies of human information processing to the various types of cognitive activity.

It is important to note that human information processing takes place in a closed loop system. Information does not flow only from receptors to perceptual mechanisms to response. Rather a very complex set of interactions occurs -- the function, and to some extent the structure, of the human information processing system is dynamic. The factors which determine the outcome of information processing vary from moment to moment.

Information about the state of the environment enters the nervous system through the receptors in the sense organs. The accuracy of that information depends on the quality (signal/noise ratio) of the displays. The probability that information enters at all depends on the extent to which the observer's attention directs the receptors to the appropriate area of the environment. Attention is sometimes automatic and unconscious and sometimes is directed consciously by decisions made by the operator which arise from recent perception and the interaction with medium term goals, short term plans, and long term memory. All these in turn are affected by newly acquired information. The detection of signals and the conscious perceptual response to the pattern recognition output are jointly determined by the quality of the incoming information and subjective biases set by the long and short term expectation of the observer as well as by the observer's understanding of payoffs associated with decisions (such as whether a miss or a false alarm is more costly).

The interpretation of the contents of perception is "diagnosis", that is, the discovery of the meaning of the perceived information. Interpretation is strongly affected by knowledge stored in long term memory and vice versa. A diagnosis leads to a decision about what action is required, to modification of plans and goals, and to modification of biases in detection and identification, and biases in the direction of attention.

As will be seen in Section 5, a rigorous interpretation of the role in error generation played by almost all the components in Figure 1 can be given. It is also easy to see how the postulated mechanisms relate to a scheme such as Rasmussen's (1983c) which describes the process control skills of the operator (See Figure 2).

Figure 1

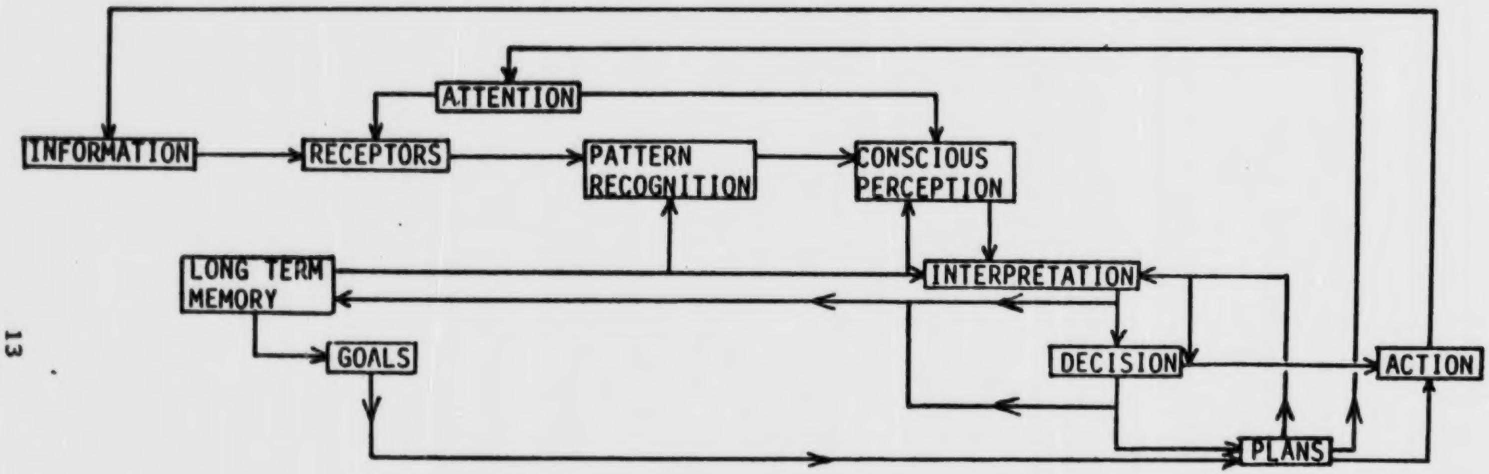

INFORMATION
RECEPTORS
ATTENTION
PATTERN RECOGNITION
CONSCIOUS PERCEPTION
LONG TERM MEMORY
INTERPRETATION
GOALS
DECISION
PLANS
ACTION

Figure 1. A flow chart of human information processing.

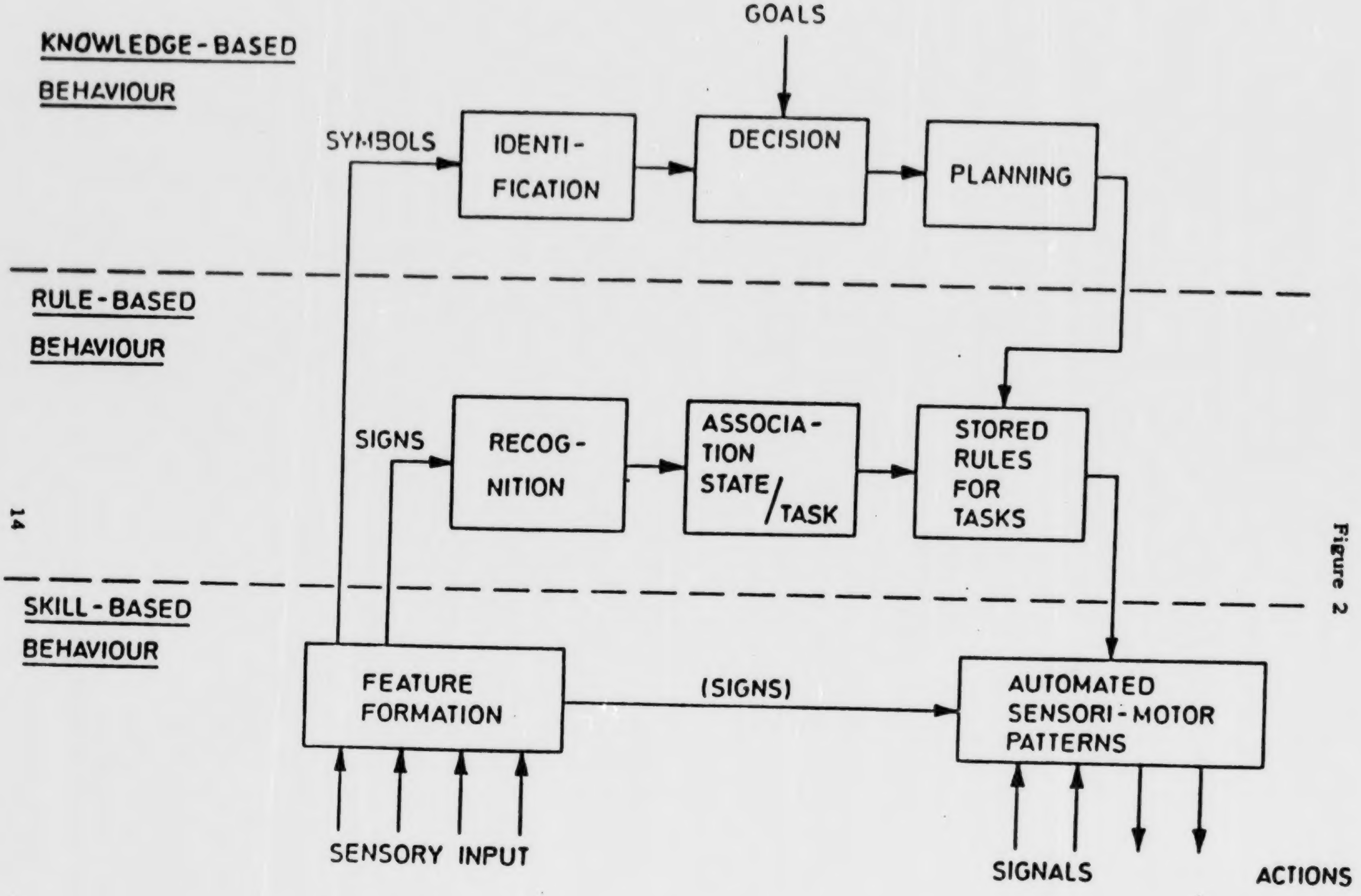

KNOWLEDGE-BASED BEHAVIOUR
RULE-BASED BEHAVIOUR
SKILL-BASED BEHAVIOUR
GOALS
SYMBOLS
IDENTI-FICATION
DECISION
PLANNING
SIGNS
RECOG-NITION
ASSOCIA-TION STATE/TASK
STORED RULES FOR TASKS
FEATURE FORMATION
(SIGNS)
AUTOMATED SENSORI-MOTOR PATTERNS
SENSORY INPUT
SIGNALS
ACTIONS

Figure 2. Process control skills of the operator. (Adapted from Rasmussen, 1983c.)

A flow chart such as Figure 1 also directs the choice of literature which is reviewed in the next section of this report. What is required is to identify the properties of the mechanisms involved at each stage -- detection, decision, attention, etc. Where possible we need to identify quantitative models for each stage. It is also desirable to discover characteristics of each stage which will limit the efficiency with which its task is performed. What are the limits on attention, on memory, on decision making which will lead to less than optimal performance?

Consider, for example, the events during the accident at Three Mile Island in 1979. (Kemeny, 1979) The operators failed to notice lamps which indicated that auxiliary feedwater isolation valves were closed. Was this due to a failure of attention (a failure to look in the appropriate direction)? Was it due to a failure to respond to the signals due to poor S/N ratio or lack of expectation that they could be lit? Was it due to a belief that the illuminated signals were themselves faulty and that the valves could not possibly be closed? Why did the operators believe for two hours that they had an overfull pressurizer rather than a loss of coolant? Was it merely because of poor instrumentation? Was it because they had to deduce the level of coolant in the reactor pressure vessel from temperature and pressure readings and failed to reason correctly? Was it because an early adoption of the high inventory hypothesis precluded them from thinking of alternative hypotheses (so-called "cognitive lockup"), or because of a failure to combine evidence from different sources optimally? Was it because the operators' long term memory contained a mental model which assumed that the light indicating the PORV status was driven by the valve stem position when in fact it was driven by the flow of current to the solenoid, and that the model was so firm that it dominated diagnostic behaviour, planning, goal formation and action?

At least some information is available from research into human information processing on all the above topics. Some is from laboratory work, other from field studies. In some cases the empirical data can be supported by or subsumed under a quantitative analytic model. But before passing on to the literature review in Section 3, it is most important to emphasize a universal characteristic of all human information processing, one particularly important in cognitive activity.

All human information processing is an interplay of the properties of incoming signals on the one hand and subjective properties of the observer on the other.

Signals can be defined by S/N ratio, and the value of S/N will in the end limit what can be decided on the basis of the signal even under optimal decision criteria. However, only if the observer adopts optimal decision criteria will the best use of the information be made. Subjective probability (expectation) and subjective utility fluctuate from moment to moment, and with them the probability of correct or erroneous judgement. Cognitive activity is dynamic and cognitive error rates need not be constant even in the face of constant displays. Figure 1 is a schematic flow chart, which shows some of the functions involved in human information processing. Three closed loops should be noted. First, the information which is available at the operator's sense organs depends on the actions which he takes; the actions themselves depend on the way in which the information is processed. Second, attention both affects which information is processed and is also diverted by the outcome of information processing. Third, past experience which is stored in long-term memory affects many aspects of information processing. It can bias attention, pattern recognition and perception by altering subjective expectations. It is the store of the operator's goals and plans. Unless the operator reads operating procedures each time the latter are used, it is his record of what actions are required by different situations. The contents of long term memory in turn depend on what information is processed and how it is processed, and hence, directly or indirectly, on what actions are taken.

3.0 LITERATURE REVIEW

3.1 The Scope Of The Review

The literature reviewed in this section is categorized as follows. First papers on the nature of the human as an information processor and decision maker are examined. Then papers dealing with the nature of human error are reviewed; in the early stages of research such conceptual studies are important. Models of error are then discussed; these include conceptual and quantitative models. The various taxonomies used to classify human error are described. Data collection methods, including significant event analysis, experimental studies, and verbal protocols are described. The latter are broken down into nuclear power plant simulator studies, other complex system simulator studies and laboratory experiments using artificial process control tasks. The relevance of fault diagnosis literature is discussed. Studies involving design to reduce error are reviewed.

This review is intended to be an organized and briefly annotated bibliography. Assessment of the literature in relation to the problem at hand will follow in the body of the report. (Also see a recent review by Moray (1985).)

3.2 The Nature of the Human Operator as an Information Processor and Decision Maker

Research on dynamic memory, decision making, mental models, and fault detection have special relevance to the study of the cognitive behaviour of nuclear power plant operators.

3.2.1 Dynamic Memory

Most academic psychological research on memory has been concerned with static memory: memory for words, syllables, numbers or pictures which are presented in lists and then recalled. The results are that under normal conditions people remember about 7 items, regardless of their nature. Comparatively little work has been done on dynamic or running memory, in which the observer must keep track of as much information as possible, when signals arrive in a continuous stream with no well defined interval for recall. Such studies as do exist suggest that dynamic working memory is much less than the "magic number 7, plus or minus 2" which is the traditional "memory span".

Work on dynamic memory suggests that observers viewing a time series do not carry more than about three items in their memory buffer, and that they probably cannot carry more. The reason for the reduction from the traditional span of seven or eight items is unclear. Kay (1950) and Mackworth (1959) both performed experiments in which the observer viewed a series of signals, each requiring a response in sequence, but with a delay. In Kay's experiment one of a row of lights would come on, then go out, another would come on, and the observer was then required to press the key corresponding not to the currently lit light, but to the previous one. Observers found it extremely difficult to handle delays of more than one signal, and impossible to handle a delay of more than two. Results were not dependent upon number of lights per row. Mackworth's results were similar.

Baker (1963) reported that when observers were detecting a series of signals which arrived at random intervals and trying to predict when the next would arrive, they appeared only to make use of the last one or two intervals to predict the next. Rouse (1973a) asked observers to predict the next value of a graphically displayed limited bandwidth zero mean Gaussian variable. He found that his observers appeared to use the current value, and the first and second differences taken over the last three signals, to predict the next. But the weightings they applied to the three signals seemed to be derived from a much longer sample, perhaps 10 to 13 items. Zeitlin and Finkleman (1975) also found a running span of three items. And Kvalseth found that even when provided with preview humans did not make use of more than one item to predict a series (Kvalseth, 1978).

The restriction to a running span of three items may only apply when the time series is "random" and "meaningless". Air traffic controllers, for example, seem to be able to recall more information from their dynamic displays. This may be due to the fact that they have, in a sense, generated the information by their commands to the aircraft rather than merely observing the displayed information. It may also be due to some form of mnemonic encoding (Yates, 1966). However even air traffic controllers rapidly reduce the amount of information remembered per aircraft as the number of aircraft under their control increases (Sperandio, 1971, 1978). It is particularly difficult to decide to what extent the limitations which appear in laboratory tasks with meaningless material apply directly to real life situations in which the operator may have, as it were, constructed the form and content of the message to be remembered. However, if one wishes to design a system in which there is the least likelihood of human error due to dynamic information being lost, a safe conclusion seems to be to minimize the amount of short term memory required. This conclusion is supported both by experimental work on the memories of witnesses in legal cases, (Loftus, 1979) and by anecdotal evidence from accident investigations, (Rolt, 1978). Bainbridge (1974, 1981a,b) on the other hand emphasizes the efficiency of memory in industrial tasks.

Some interesting theoretical papers suggest that surprisingly small memories may be adequate for dynamic decision making. Baxa and Nolte (1972), Gonzales and Hovington (1977), Mullis and Roberts (1968), and Taylor (1975) all discuss the question of how much memory is as good as an infinite memory; Bechtel and Amos (1975) discuss the tradeoff between the number of dimensions to be examined and the number of samples per dimension, and its effect on the miss and false alarm rates for a detection task on a variable of high dimensionality. Those papers are not concerned with how humans behave, but with the problem of optimal decision making based on limited samples. The rather surprising conclusion is that a memory buffer with seven or eight entries can perform almost as well as a buffer of indefinite size if the task is to accumulate noisy evidence for decision making. Below about four or five entires performance is very badly degraded. It seems therefore possible that even if one agrees with Rouse (1973a) that when humans observe sequences of signals they extract weighing functions from longer sequences but only weight the last two or three items, then the operator would be found less badly handicapped than the empirical data on dynamic span at first suggest.

An important exception to the very limited properties of memory outlined above is recognition memory. Experiments by Shepard (1967) showed that an observer matching a new stimulus against examples can recognize one seen earlier with an accuracy which is orders of magnitude more efficient than when the same observer tries to recall an item from memory. If it is essential that memory is used in a task, every effort should be made to ensure that recognition, rather than recall can be used. This is a very well established result.

3.2.2 Decision Making Ability

At least four major reviews, in addition to Peterson and Beach (1964), are unanimous in their judgement on the inefficiency of humans as statistical decision makers. These are Dale (1968), Hopf-Weichel, Lucaccini, Saleh, and Freedy (1979), Sage (1981), and Schrenk (1969).

Most writers agree that humans can estimate the mean of a series of numbers, and are good at estimating proportions, although there is usually a tendency to underestimate high proportions and to overestimate low ones. They are also good at understanding and using probability statements which are not based on frequency, such as "There is a .1 probability that it will rain today". But apart from that, man's record as an intuitive statistician is poor. Sage (1981) gives no less than 27 aspects of decision making which cause some sort of bias away from the optimal in the human statistical decision maker. Vaughan and Mavor (1972) state that man

> "is slow to initiate action and conservative in his estimates of highly probable situations, and when he does act or accept a diagnosis, he is reluctant to change an established plan or a situational estimate when the available data indicate that he should. He is generally a poor diagnostician and does not learn by mere exposure to complex tasks but only when specific relationships that make up the complexity are explained. He is not particularly inventive and tends to adopt the first solution he finds. He finds it difficultto use more than one criterion at a time in evaluating actions and tends only to identify criteria that reflect favorably on the action he is developing. He tends to use only concrete, high confidence facts in his planning and prefers to ignore ambiguous or partial data rather than attempt to interpret them." (Vaughan and Mavor, 1972, p. 274)

Although the above was written in the context of medical diagnosis and decision making, those familiar with the events at Three Mile Island (Kemeny, 1979; Mason, 1979; Rubenstein, 1979) will recognize the description.

As the complexity of the system which humans must control grows by orders of magnitude, the problems of diagnosis become less and less like fault tracing in a radio, for instance, and more and more like medical diagnosis, because of the degree of ambiguity and the looseness with which symptoms are related to a specific cause. Although the papers by Dale (1968) and Schrenk (1969) are now nearly 15 years old they still remain excellent summaries, and suggest the following conclusions, which are supported by the more recent paper by Sage (1981).

> Humans are bad at estimating variance, and usually underestimate it; they are particularly bad at combining evidence to update probability estimates, and are usually conservative, refusing to change as much as is appropriate. They request more evidence than is necessary before coming to a decision. They are both strongly biased by events early in a series, and strongly dependent, once into a series, only on the most recent data. They form hypotheses early, and then try to confirm them rather than testing them and they do not consider enough hypotheses. They believe small samples to be more typical of population than is warranted. They tend to overestimate the

probability of outcomes if the outcomes are favourable, and underestimate the probabilities of outcomes if the outcomes are unfavourable.

In addition, humans seem unwilling to accept statistical rather than causal relations, and will try with grim determination to impose a causal account of variability on stochastic events (Gaines, 1976).

Generally the Bayesian model has been taken as the norm against which to test human behaviour, and almost without exception the human is suboptimal. The Bayesian model assumes a completely rational observer who makes the most logical use of evidence to estimate probabilities by revising estimates in the light of new evidence. Human beings do not revise their estimates of probability as much as the evidence warrants when presented with a series of pieces of evidence. There are alternative explanations for some of the experiments. It has been suggested that the results of Tversky and Kahneman (1971) showing that human beings see small samples as more typical of the population than is really the case, and their experiments on representativeness, may not show non-Bayesian behaviour, but rather Bayesian behaviour given incorrect estimates of the odds. There are also some violations of the generally conservative attitude to the combination of probabilities, but when exceptions are found, they do not lead to optimal behaviour but to behaviour which is suboptimal in a different way.

A more hopeful view of humans as processors of uncertain information is provided by studies such as those of Levine and Samet (1973) and Levine, Samet and Brahlek (1975) which show that humans are able to detect and make use of the differential diagnostic value of information and will choose to ask for more information from sources of good quality information.

Other aspects of decision making, especially in connection with predicting the future values of a time series on the basis of present and past evidence, are to be found in the work of Curry and Nagel (1976), Phatak and Bekey (1969) and Rouse (1973a, 1973b, 1977). The results are in line with the studies already mentioned. For example, Curry and Nagel presented a display in which a spot of light moved down a screen in a series of steps toward a target at the bottom of the screen. They asked observers to decide on which side of the target the spot would pass. Many of the observers were very resistant to changing their hypotheses, a finding in agreement with Rouse's experiments. Furthermore, the observers did not predict as far into the future as the evidence warranted.

There are many tasks which, on the above evidence about human decision making, humans should not be able to do, but which they do rather well. For example, work such as that of Coeterier (1971), Sperandio (1971, 1978) and Soede and Coeterier (1971) on the behaviour of air traffic controllers, suggests that memory is much better than three items. And the work of Bainbridge (1974, 1978a, 1978b, 1981a) and Bainbridge, Beishon, Hemming, and Splaine (1974) all indicate that process control operators have far more information available to them about the complex dynamic systems they control than laboratory experiments would lead one to believe possible.

Two differences seem critical. The first is that in field studies and daily "real life" control, the controllers bring about the situation which they are trying to remember. Although this effect has not been studied in detail, the evidence suggests that people find it easier to remember many variables when they have set the variables to their value rather than merely observing them. Sperandio (1971, 1978) however indicates some of the limits on even this skill. As the number of aircraft being controlled increases, the number of items of information retained about each decreases, and the controller progressively constrains the behaviour of the aircraft to simplify his task, thus reducing the memory load.

The second point is that all the tasks where behaviour is found to be surprisingly close to optimal, the operators or observers are very highly practised. Optimal behaviour, well described in optimal control theory and optimal estimation theory by complicated matrix manipulations which process information from several sources in parallel, is not seen until the behaviour becomes almost automatic; the operators are not thinking consciously about what they are doing. By contrast, all the laboratory studies on "humans as intuitive statisticians" use relatively unpractised people who make their statistical decisions consciously. Even more important, perhaps, is that most subjects in the experimental laboratory work with systems which are all "false front" with nothing behind. Real world systems on the other hand, are coherent and logical. Only for such as these latter can the human operator establish a mental model of the plan and gain an improved ability to deal with the system properly.

It is as a conscious statistician and decision maker, reflecting and working in a knowledge-based mode of operation, that humans fail. As highly practised, automatic operators, using skills which require no conscious reflection, we become optimal. Skill-based behaviour is far more effective and optimal than problem-solving behaviour. If one wishes to see optimal behaviour, prolonged practice to the point where the behaviour is automatic is required.

3.2.3 Response to Faults and Emergencies

Since the main role of the human supervisory controller is to detect and diagnose failures, a crucial problem is to ensure that the operator has appropriate models for different tasks. It seems likely that the efficiency of the operator in emergencies will depend to a great extent on the correct "cognitive coupling" (Green, 1980) between the operator and the task. How volatile such coupling may be is shown by an experiment by Shiff.

Shiff (1983) trained operators to control a computer simulation of a pulp and paper mill. The operators practised for several hours using manual control. He then divided them into four groups, each of which operated the mill for a further four runs in different control modes. These were manual control, computer feedback as to whether the action taken was good, computer advice as to what the computer thought was a good strategy, or full automatic supervisory control, in which the operator merely watched the computer open and close valves, start pumps, and so on. On the final run the computer "broke down", and the operator was required to take the system back under normal manual control. Despite the fact that all operators had received substantial initial training with manual control, those who had subsequently had a short period of supervisory control were very slow and inefficient at bringing the plant under manual control, and their trajectory through the problem space was far less efficient than that of those who had continued to use manual control.

The detection of abnormal and fault conditions in the field setting of a large refinery, power station or aircraft is a far more complex problem than most laboratory situations. De Keyser (1980) has drawn attention to the complexity of the skills used by process control and power station operators in such diagnosis, including the fact that they make a great deal of use of what she calls "informal information" such as vibrations, smell, and so on in addition to the "formal information" from the control panel. She warns that the gradual removal of operators from physical contact with the process may impoverish their diagnostic environment despite the more sophisticated displays.

This is particularly important if operators do not try to track all the state variables, but make use of a subset of them together with an internal mental model which, if properly updated, is sufficient to allow them to control the system. It is often assumed that practice builds such models, but their nature and properties are still quite obscure. An internal model should, if one wanted evidence for the reality of mental models, allow the operator to exercise control in some situations which have never been encountered before, providing that the situation is logically entailed as a possible consequence of the model. Moreover, such conditions should be sufficiently different from those encountered in training that simple extrapolation in the sense of "generalization", as that term is used in learning theory, is implausible. The idea of a model is that computation based on the acquired knowledge base is needed, and possible, even though the situation is new. Furthermore, the idea of a model usually allows the use of analogy, so that knowledge from other but "similar" experiences will be accounted relevant. Although the concept of an internal or mental model is so intangible, great interest has recently been shown in the idea in many areas of psychology (Gentner and Stevens, 1983; Johnson-Laird, 1983). Extensive discussions of models can be found also in symposia edited by Sheridan and Johannsen (1976) by Rasmussen and Rouse (1981), and in several papers in Edwards and Lees (1974).

3.2.4 Fault Detection and Diagnosis, and Human Error

There is a class of research which at first sight appears relevant to the question being examined in this report. Although we do not believe it is in fact relevant, we feel that it is necessary to justify its omission.

Rouse (1981), Gai and Curry (1976), Wewerinke (1981, 1983), Wohl (1982) and several others have all examined how long it takes an observer to notice a "fault" in laboratory settings. Beare et al (1983) and Kozinsky et al (1984) have asked the same question about operators in BWR (Boiling Water Reactor) and PWR (Pressurized Water Reactor) simulators. Rouse asked operators to trouble shoot simulated logic networks and simulated or real airplane engines. Wewerinke, and Gai and Curry, asked observers to look for a change in the mean or variance of a bandlimited Gaussian noise. Wohl examined trouble shooting in defective radar sets. In each case the operators or observers continued to work until the fault was found. Hence there were no measureable errors by the operators. What was measured was how long it took to find the correct solution, to detect the change in the statistical properties of the display, or to initiate the correct remedial action to a NPP transient.

Now it is possible that errors did occur. Perhaps the reason why some of Wohl's technicians required several hours to complete the diagnosis was that they performed incorrect tests. Gai and Curry's observers may perhaps have misjudged the magnitude of the function which they observed. But in none of the studies are there data which directly throw light on the occurrence of error, only on the time taken to find the correct solution. If there were a standard time to perform each task, then any case in which the operator required more than the standard time could be interpreted as an "error". But we then could be faced with the peculiar assertion that an operator who carried out a fault diagnosis procedure in every respect correctly but rather slowly was making an "error", whereas someone who made several unorthodox moves but completed the diagnosis earlier than the standard was "correct" even if in fact two tests were incorrectly used but happened to cancel each other out.

Studies on system faults are not studies of human error. They are of inherent interest and of great importance. For example, Beare et al (1983) in particular is thought provoking in the context of training, design, simulator practice, and safety in NPP operation, but it does not speak directly to the aims of the present report. For that reason such studies will not be discussed further here. Those interested in more detail should consult the original papers or the review by Moray (1985), or, (when it is published), a report from the Electrical Power Research Institute by Hanneman, Spurgin and Lukic which is presently in draft form.

3.3 Mental Models And The Prediction of Error

3.3.1 Characteristics of Mental Models

One finds statements in research papers to the effect that operators control not the real system, but their "mental model" of it; that displays should support the operator's "mental model", and so on. Surprisingly, the concept had wide support among engineers and engineering psychologists long before it became current in more traditional psychology. It is clear that the study of humans working with very complex systems forced researchers to consider a concept which was not obviously necessary to psychologists who performed experiments using very simple laboratory paradigms. Recently there has been a great surge of interest in the idea of "mental models" among psychologists (Gentner and Stevens, 1983; Johnson-Laird, 1983), and it is clear now that the concept will be regarded in future as central to an understanding of human information processing.

A model in the sense used in this section refers to a less than completely accurate representation in one medium of a system that exists in another. In the case of NPP operators we are referring to the representation in the brain of the structure and properties of the NPP. During training and operations an operator acquires knowledge about the plant in several ways. Firstly there is "academic" knowledge, a list of facts, properties, formulae, etc., expressed in words, numbers, graphs and pictures. Secondly, there is "applied" knowledge, a feel for the dynamics, kinetics, etc. of the system which is built up by operating a simulator or the real plant and by acquiring direct "hands on" experience with the physical entities which make up the plant. Thirdly there is operational skill, the ability to act correctly even if the operator is unable to express in words why he takes the appropriate action. Somehow, it is assumed, these three sorts of learning build a representation of the system in the operator's brain. It may enable him to talk about the state of the system to himself or to others; it may lead to visual images in his mind's eye of mass balances, flows, energy fluxes, as well as pumps and valves. This representation is the "mental model".

Almost certainly it is not a veridical representation of the real plant. In many cases the operator will diagnose correctly plant states which he has never seen before either in the simulator or real NPP when the states occur for the first time, because his model will allow him to extrapolate from the old to the new situation. Equally, there will be cases in which no extrapolation will be possible, and in which his model will be inadequate. If training has been effective, it is reasonable to expect a wide range of procedures to be available to him: that is, the model will include a substantial amount of "rule based" behaviour. But the concept of such a model includes far more than the idea that the operators' knowledge consists of a set of factual relations of the form:

IF (situation A) THEN (action B.)

Rather, the contents of a procedure will become related to all other meanings of the words which embody it, and to other similar procedures to a greater or lesser extent. Training and operation produce a mental model which differs in many ways from the intentions of the trainer.

Because of this, when the model is used by the operator particularly in "knowledge-based" ways, for assessment, diagnosis, and planning, it can generate cognitive errors.

The operator experiences the model as the intuition he feels about the meaning of status information. As such he is unaware of the extent to which the model is or is not a veridical representation of the plant. A classic example is the certainty of the operators at TMI-2 that there was no loss of cooling inventory because pressurizer levels were always displayed as above normal. Because of the contents and relationships in their mental models it was effectively impossible for them to think about the plant in certain ways, until evidence overwhelmingly suggested that their models must be wrong. Operators will interpret evidence and choose actions in terms of their mental models. Only in extremely exceptional circumstances will "logic", "reason" and "rational thought" be substituted for "running the mental model".

In some kinds of behaviour there is strong evidence for a mental model. In learning to control a complex or possibly unstable machine like a helicopter, success comes with a progressive reduction of latency and a high level of skill is largely manifested in anticipatory behaviour. Such anticipation must be based on an organized reservoir of information about the couplings of the various degrees of freedom and of the effect of derivatives of state variables. These are the mental model which permits easy efficient operation.

It is clear that the idea of the mental model has tremendous potential for explaining and predicting both effective and erroneous behaviour. The concept is far more than a heuristic convenience. Unfortunately there is little or nothing known about the nature, content, acquisition and properties of operator mental models. Some of the best description of operators' models are those given by Bainbridge (1974, 1978, 1981(a)). While suggestive, they cannot be used to predict error in specific NPP situations, and even the sophisticated discussions in Gentner and Stevens (1983) and Johnson-Laird (1983) cannot at present be used.

3.3.2 Methods for Analysis of Mental Models

Verbal reports, in which operators think aloud while performing operations, or immediately after, are the most intuitively obvious ways to collect data on operators' mental models. A review by Bainbridge (1979) covers system state/action state diagrams, questionarie, interview, static simulation, and verbal protocol techniques. She points out the difficulties in probing covert mental activities through verbal reports. What people say that they do is not necessarily what they actually do. The few data that exist for the correlation between observed behaviour and reports emphasize the amount of distortion that occurs in verbal reports. These techniques assume that knowledge exists in a verbal form, that people are aware of their mental processes, and that mental activity always affects behaviour. The problem with off-line reports is the well documented distortion that occurs in memory. Report data are also open to the subjective interpretation of the experimenter who infers the mental model of the subject by processing it through his/her own mental models (preconceptions, hypotheses). Eriksen and Simon (1984) discuss the nature and use of protocol analysis in laboratory investigation.

Landerweerd (1979) tried to teach chemical plant trainees two different types of internal representational models. Verbal, classroom instruction was hypothesized to cause the trainees to form an abstract, internal representation of the chemical process while training on a simulator was intended to cause another group of trainees to form a visual, concrete representation (mental image) of the same process. Fault detection, diagnosis and correction performance measures showed that the first group was better at fault correction while the second (mental image) group was better at searching the relevant information and diagnosing the fault.

Bisseret (1971) analyzed the mental models of air traffic controllers by interviewing them and asking them to describe their general reasoning process while solving an ATC problem. He then gave them a specific ATC problem and repeated the interview process. Finally, he posed a problem to them and gave them no further information. Their internal model was inferred from the questions that the controllers asked. He concluded that the main mental task of the controllers was one of categorization and that a flow chart depicting this process fitted actual diagnostic data well.

Veldhuyzen and Stassen (1977) incorporated the concept of the internal model into a control theory model of a helmsman steering a ship. A computer simulation of the control model could predict actual performance data reasonably well with some minor modifications.

Viewing the plant at the wrong level of abstraction has broad implications for how errors are committed and when they are detected. More specifically, people tend to seek information that confirms their expectations. If an anticipated event fails to occur, then an error will be self-detected. Often, in errors of diagnosis, the wrong plan is unintentionally activated, and since humans tend to seek out confirming evidence that the plan has been carried through, a whole sequence of actions may have been executed before the misdiagnosis is discovered.

It is quite certain that an understanding of these topics, and of the effect of mismatches between the mental model and the real plant, would provide a powerful tool in error prediction. If we accept the model of error which suggests that errors arise because of the similarity in the "path" followed by information to another route (Senders, 1983; Norman, 1983b), then when the mental model contains knowledge which is similar to but not exactly the same as the real system, it will produce an easy opportunity for processing to be "diverted" down an incorrect "path".

At present the idea of the operator's mental model, and the degree to which it fails to match the properties of the real system must remain as a qualitative heuristic. We know little about how to reveal models other than by the use of protocol analysis with all its problems. It seems likely that the generic properties of mental models differ from operator to operator, and also are plant-specific. Task-specific, operator-specific models may need to be examined. At present no method for this exists.

It has been suggested that if designers and trainers aimed to produce an optimal match between the system and the operator's internal model or models, this would facilitate diagnostic behaviour. However there is also the problem of the operator's decision-making psychology, which as we saw tends to produce "cognitive tunnel vision", i.e. a tendency to stick rigidly to the first hypothesis which he finds attractive rather than exploring alternatives. Moray (1981) has pointed out that such behaviour can, in some cases, be the correct strategy, because when a fault is detected, its cause is more likely to be in subsystems which are closely coupled to the faulty subsystem than in a subsystem only distantly coupled. Hence a rational search strategy is to keep testing only tightly coupled components.

On the other hand one would expect the opposite strategy when the operator finds only normal readings. It should be possible to rely on the coupling between the subsystem and so use a relatively sparse sampling strategy. Indeed this must be the way in which operators scan very large displays, since there is not time for them to look at every display. Iosif (1968, 1969a,b) has reported data which support these suggestions.

It is probably fair to say that while the acceptance of the idea of an internal model is widespread, little is known about how to use the concept for design or analysis other than in a post hoc way, with the exception of Optimal Estimation Theory. The discussion by Bainbridge (1981b) is particularly important. There is, moreover, at least one alternative explanation of the human observer's detection and diagnosis of failures, and that is pattern matching. It is not inconceivable that over a long period of practice the observer comes to know what certain abnormal patterns look like, and has associated with each pattern an appropriate set of actions. Gonzales and Hovington (1977) have reported how powerful computer pattern matching can be in fault diagnosis, and Chu, Steeb and Freedy (1980) have shown that adaptive pattern matching can be a model for action sequences. It might be hard to distinguish such a system from the currently popular "internal model". Probably the single strongest argument against pattern matching as the way in which the observer performs all but a handful of diagnoses is the extreme rarity of faults in large automatic systems. It seems very unlikely that enough experience could be acquired to build up the range of templates needed to use pattern matching effectively. It is not however out of the question for some situations, and for some training programs.

3.4 Concepts and Models

3.4.1 The Nature of Error

The study of error for its own sake rather than as a black-box measure of performance has only recently begun. Senders (1983) points out that most psychologists ignore error as a psychological process and use it only as an index, when in fact, errors are psychological events worthy of investigation in their own name. This is particularly true with increasing automation, where fewer people have greater responsibility and the consequences of error have grown accordingly. Senders concludes that "it is errors that destroy us, yet all we have done is to count them and calculate probabilities. We know little about how ofter errors occur, what they are or will be".

Norman (1980) also emphasizes the usefulness of studying error. Using Tenerife and Three Mile Island as examples, he shows that accidents have multiple causes and can include cognitive, physical and social factors. Because of this accidents are difficult to classify. In comparison errors are more numerous and more easily classified.

3.4.2 Concepts of Error.

Definitions of what constitutes an error vary. Swain and Gutmann (1982) state that only errors which lead to important negative outcomes are worthy of attention. However, Adams (1982) argues that, if human errors are to be used to define human reliability probabilities in the same way machine failures define machine reliability, then all errors should be looked at, not only those that generate system failures. After all, with equipment reliability, component reliability determines system performance not vice versa. Adams also notes the problem of definition when an error is self-corrected. Sheridan (1983) questions whether some actions may be incorrectly recorded as error when they are not in fact, and offers the example of performing a step out of sequence when the sequence doesn't matter. Related to this is Sheridan's contention that errors should perhaps be described in terms of degree rather than using an all or none description.

Numerous causes of error have been postulated. Norman (1983b) postulated that Freudian slips account for some errors, while others are due to deficits in training or equipment design. Senders (1983) postulates that while some errors are "caused", others are not and are random. He refers to laboratory data which show that errors appear to occur with a constant probability per opportunity, and thus cannot be predicted in time (Sellen, 1983).

The problem of mechanistic rather than psychological approaches to the study of error and reliability is discussed by a number of authors. Adams (1982) states that while the idea of human reliability is a sound one, methodological problems make it difficult to express human reliability. The difficulty of measuring human reliability is related to the problem of determining behavioural units or taxonomies, and the difficulties, noted earlier, of defining when an error has occurred. The psychological approach considers such factors as learning, motivation, purposeful behaviour etc. which affect behaviour only in humans but not in machines.

Hollnagel (1983), like Adams, argues for a psychological rather than an engineering analysis of the human part of the human-machine system. The latter, he claims, can only address those aspects of the human that are machine-like. With the increasing complexity of systems, the inadequacy of this approach becomes clearer. The psychological approach, on the other hand, makes clear the fact that there are a number of areas in which we lack important information, for example on how performance is shaped, how strategies are formed, and how mistakes in decisions occur.

Sheridan (1980) discusses the differences between machine error and human error illustrating some of the difficulties in estimating human reliability. The following points are worth mentioning:

1) There are many different theories that may be used to account for the cause of human error, unlike machine error which may be explained in terms of the laws of physics.

2) There is no single agreed upon taxonomy of error and hence data collection may not be as objective as desired.

3) People can recover their errors and so a true measure of human error rate is difficult to obtain.

4) More than one operator is likely to be working at one time and the social interactions may affect human error rates in complex ways.

Models of human cognitive performance can be categorized on a number of dimensions. There exist qualitative or conceptual models such as Norman's (1983b) verbal analytic model of human error, and quantitative models such as HOS -- Human Operator Simulator (Lane et al, 1980), a collection of computer programs, which can be used to predict an operator's performance, including decision and execution times.

Another way of classifying models is psychological vs. engineering based. Psychological models are often bottom-up in that they concern themselves with basic behaviour, such as information processing, decision making, movement etc., which are systematically organized to describe task-oriented behaviour. Engineering models on the other hand are top-down in that they start with an description of the human-machine system as a whole, and attempt to model the human component in modular fashion at the task or function level.

A third dimension of models is the scope of description. Only a few are overall models, for example HOS, which attempts to describe the whole of human performance. Most are micro-models which look at one aspect of performance such as decision-making or information processing.

In the following section conceptual models will be discussed first followed by quantitative models.

3.4.3 Conceptual Models

Reason (1977) presents a model of skilled behaviour which is behaviour rather than task oriented. The model contains information inputs and outputs, an intention system and store, and an action system and store (see Figure 3). Reason focuses on how everyday errors are generated; he calls these 'actions not as planned'.

The intention system formulates the plans and initiates action to execute them and tests their progress. This last provides a feedback mechanism. The intention system is said to be embodied in the current contents of consciousness. Reason suggests that more than one plan may be active at one time. He bases this on evidence that under some circumstances, particularly where humans are highly practised or the task has a high degree of input-output compatibility, parallel processing of information seems to occur. The activation of more than one plan can, of course, lead to error.

While the intention system acts as the executive, the action system is the operations branch. It is responsible for structuring well-practised sequences of motor commands. The action system feeds back information to the intention system about how the plan is progressing. An input function allows selected signals into the intention and action system. These signals may be from the outside world or feedback about the actions themselves or the actual outcome of the action.

Storage systems are postulated which contain previous plans (intention store) and sets of actions (action store) to carry them out.

Figure 3.

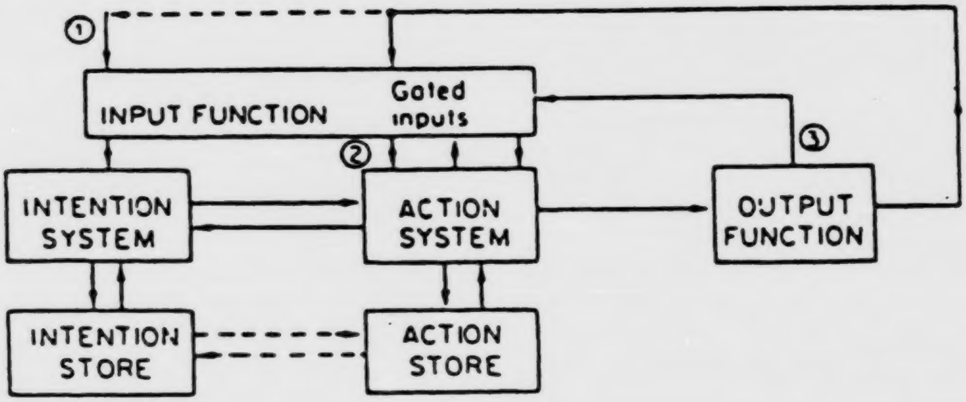

①
INPUT FUNCTION
Gated inputs
②
③
INTENTION SYSTEM
ACTION SYSTEM
OUTPUT FUNCTION
INTENTION STORE
ACTION STORE

Figure 3. Flow diagram showing the basic components of the composite-skill model. The ringed numbers refer to inputs from the outside world (1), feedback from the actions (2), and from the actual outcome (3). (Reason, 1977)

Reason then sets up an algorithm, based on his model, which can be used to classify errors as storage, test, discrimination, or selection failures (see Figure 4). This classification system is discussed further in the next section which deals with taxonomies.

Two other models which have the same basic structure as Reason's were proposed by Norman (1981) and Senders (1983). Norman considers actions to be sequences of activation-triggering schemata. The schema is defined as an organized body of knowledge which directs the flow of control of motor activity. Errors occur in the formation of intention, triggering of the schema and activation of the schema. Reason's 'input signal' is equivalent to Norman's triggering. While Reason classifies errors into four types related to the interaction between intention and action, Norman's classification of errors is much less structured. These classifications will be discussed further in the section on taxonomies.Senders (1983), like Reason and Norman, considers three sources of error production which he calls intention, perception and action. Senders classifies possible errors in four categories: omission, repetition, insertion and substitution. While Norman and Reason present models which are post hoc classifications of error, Senders' model includes prediction of error. He quotes Kollarits' (1937) data showing 70% of errors were substitution errors. With respect to substitution errors, Senders (1983) proposes a quantitative theory of how such errors come about and states a number of experimental hypotheses derived from the theory. Figure 5 illustrates several behavioural sequences. There is a domain of intentions (on the left) which result in a domain of actions (on the right). Sequence A.a and B.b have a number of common elements. Senders' hypothesis about the generation of substitution errors is as follows:

1. The probability of a substitution error will increase as the number of common elements increases.

2. As the number of steps from the last common element in two sequences to the action increase, the probability of a substitution error decreases.

3. For a given number of common elements, the probability of a substitution error will increase with a decrease in memory of the intention.

4. The addition of redundant or meaningless steps between the last common element and the action to be completed will decrease the probability of substitution errors.

Because these hypotheses are as yet untested and are unsupported by data, they remain a theory rather than a model per se. However this theory is important because, to the authors' knowledge, no other general conceptual predictive model exists currently.

Janis and Mann (1977) present a conceptual model of emergency decision making in response to disaster warning which integrates various findings from psychological research in this area. The first level of the model defines the antecedent conditions, i.e. the quality and type of information about the disaster. The second level of the model is a series of questions people ask themselves when faced with an emergency, the answers to which determine choice of action. The third level consists of the possible coping patterns which might be used to respond to an emergency, for example hypervigilance or panic. The conditions required for effective emergency decisions are discussed. Suggestions are made for ways of facilitating emergency decision making.

Figure 4.

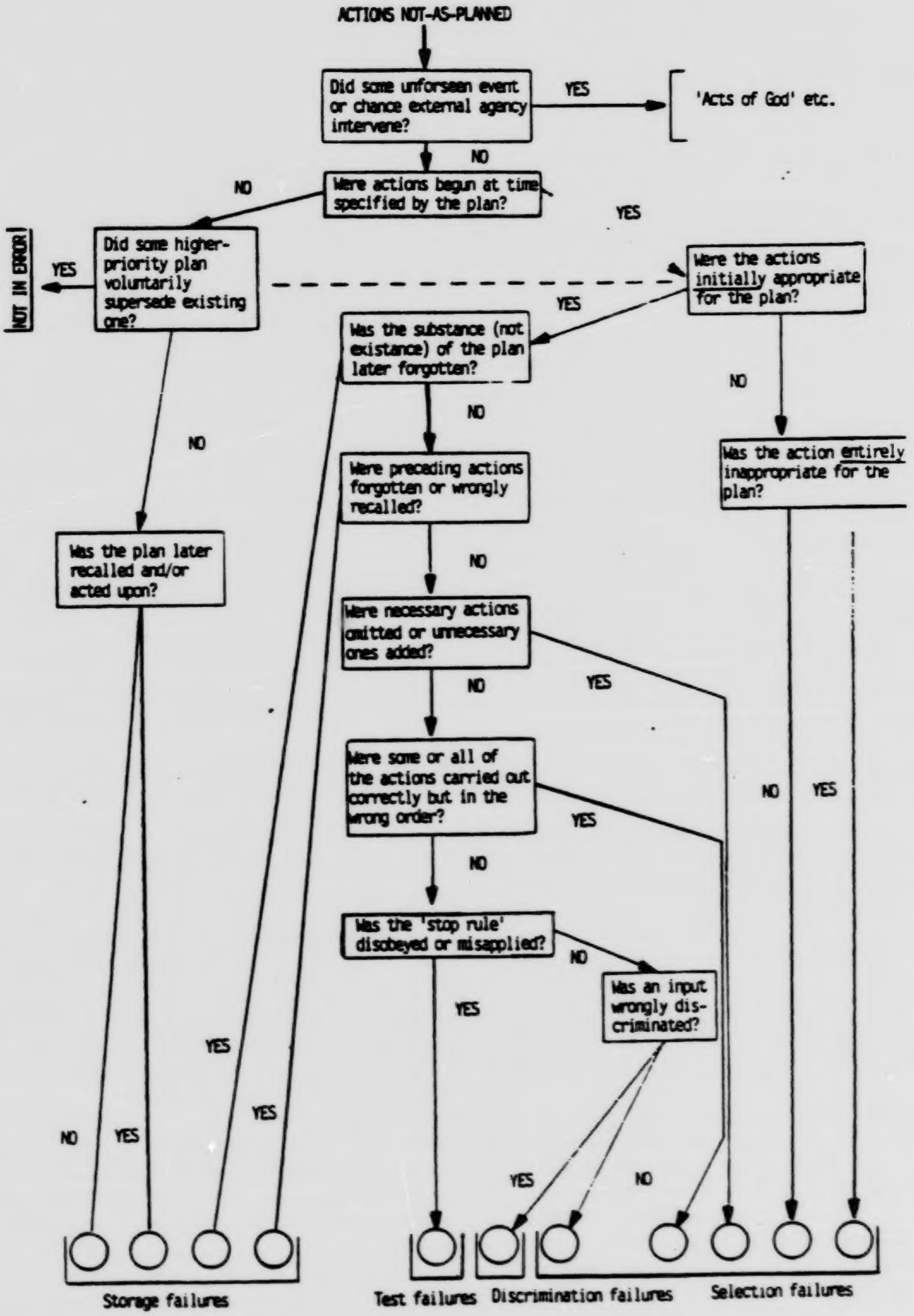
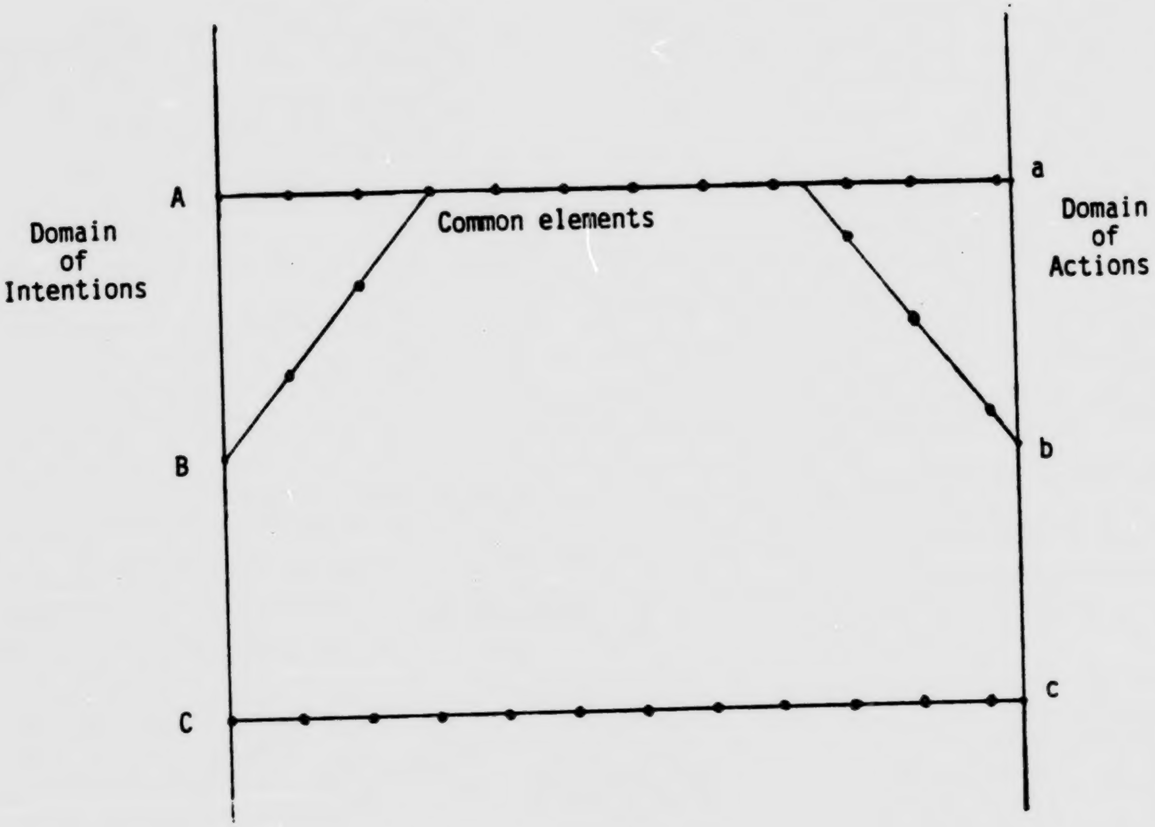

ACTIONS NOT-AS-PLANNED
Did some unforseen event or chance external agency intervene?
YES
'Acts of God' etc.
NO
Were actions begun at time specified by the plan?
NO
YES
Did some higher-priority plan voluntarily supersede existing one?
YES
NOT IN ERROR
Were the actions initially appropriate for the plan?
YES
Was the substance (not existance) of the plan later forgotten?
NO
NO
NO
Were preceding actions forgotten or wrongly recalled?
Was the action entirely inappropriate for the plan?
Was the plan later recalled and/or acted upon?
NO
Were necessary actions omitted or unnecessary ones added?
NO
YES
Were some or all of the actions carried out correctly but in the wrong order?
YES
NO
YES
NO
Was the 'stop rule' disobeyed or misapplied?
NO
Was an input wrongly discriminated?
YES
YES
YES
NO
YES
YES
NO
Storage failures
Test failures
Discrimination failures
Selection failures

Figure 4. Algorithm for classifying actions not-as-planned. (Reason, 1977)

Figure 5.



Domain
of
Intentions
A
B
C
Common elements
a
b
c
Domain
of
Actions

Figure 5. A sample of error production.

Rasmussen (1981) developed a task-oriented model of cognitive behaviour. He describes the latter as being skill-based, rule-based or knowledge-based. Skill-based behaviour is exemplified by over-practised, automatized perceptual-motor skills. Little such behaviour is seen in NPP control rooms. Rule-based behaviour is exemplified by procedure following. Procedures may need to be looked up in books or data bases, or may be recalled from memory. Their characteristic structure is pairs of operations: IF (situation) THEN (action). Knowledge-based behaviour is the use of thought, analogy, logic, etc. to understand a situation and decide on action when neither skill nor rules are available or appropriate. It is, in a word, cognitive and conscious. In rule-based situations perception can lead to action at the skill-based level: perception leads to identification, which in turn leads to a choice of procedure, which leads to action. In knowledge-based situations it is the incomplete identification which leads to the formulation of hypotheses, to tentative diagnosis, to the formulation of a goal, to the choice of procedures and to action. Cognitive errors are particularly related to knowledge-based behaviour. For a novice no behaviour is automatic; for an experienced operator, very complex behaviours can be skilled-based or automatic. Thus the type of behaviour is somewhat dependent on the operator's training and background.

A model of the human decision process in terms of skill-based, rule-based, and knowledge-based behaviour is shown in Figure 6. The importance of thinking about behaviour in this way is that it helps determine the approach to reducing errors. For example if errors occur in skill-based behaviour, it may be because the operator did not receive enough training on a specific skill, or possibly performance shaping factors such as poor motivation, fatigue, or stressful working environment are affecting behaviour. Errors in rule-based behaviour can be reduced by the use of checklists or by teaching operators the reasons for the rules. Errors in knowledge-based behaviour can be dealt with by more in-depth training of operators as to how the system functions.

Pew and Baron (1983) note that an important feature of Rasmussen's model is that it allows for short cuts. Rasmussen argues that an observer may detect a problem, collect limited data, and conclude immediately that a specific control action must be executed. Such a short cut he calls skill-based behaviour because the specific features have been experienced together frequently before and the response is more or less automatic.

Alternatively, the control room operator may detect an alarm, collect data, identify the system state and then immediately select and execute a procedure that results in an action sequence. This is called rule-base behaviour by Rasmussen and reflects a higher level short-cut.

3.4.4 Quantitative Models

Schurman and Banks (1985) have identified and reviewed nine models for predicting human error probabilities or error rates. (Their report includes a discussion of data collection methods, measurement theory, and reliability estimation.) The nine include some (such as THERP) which have already been in use for a number of years, others (such as OATS) which have had relatively slight use, and yet others (such as STAHR) which are still under development.

Figure 6.

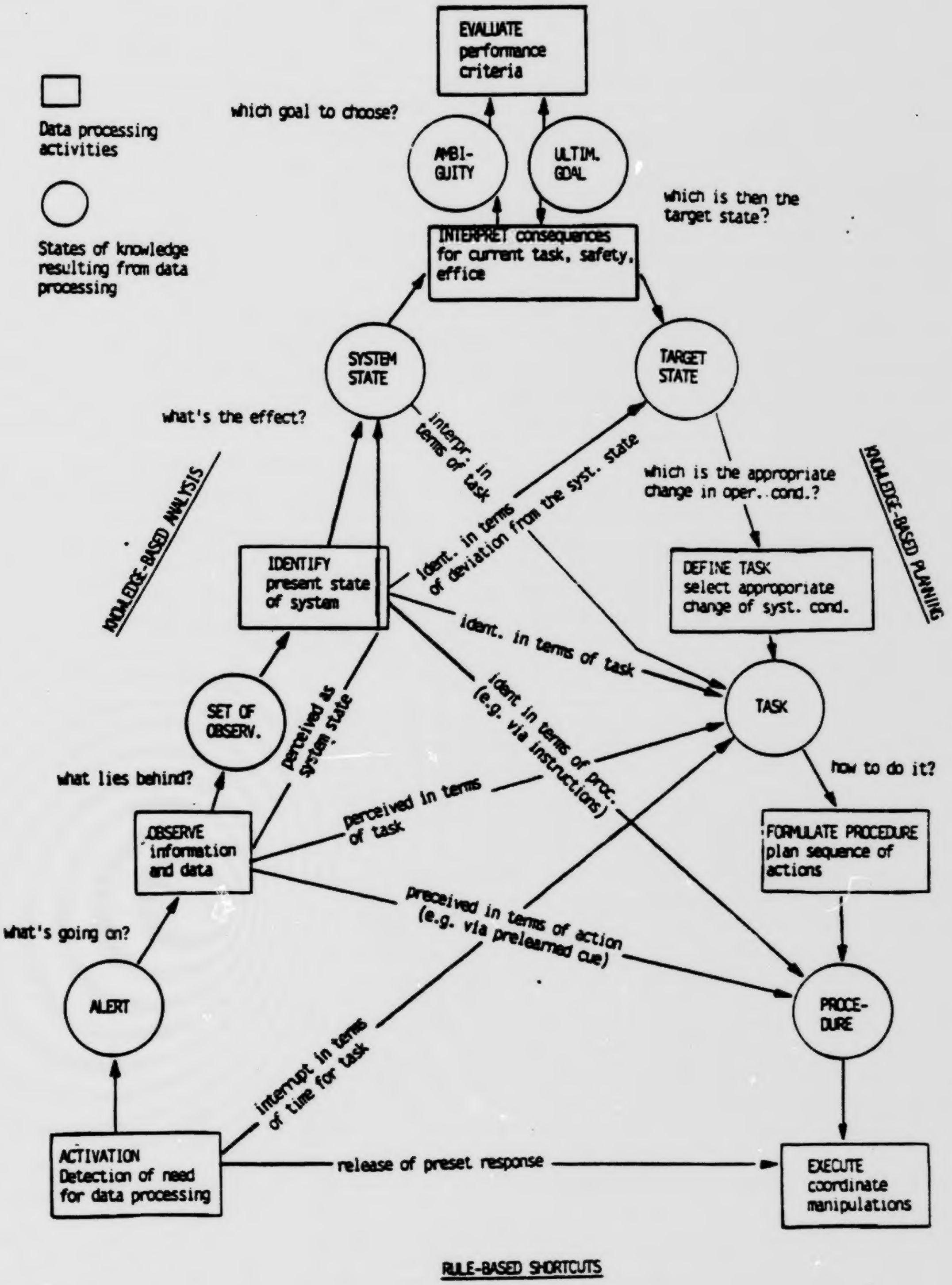

Data processing activities
States of knowledge resulting from data processing
EVALUATE performance criteria
which goal to choose?
AMBI-GUITY
ULTIM. GOAL
which is then the target state?
INTERPRET consequences for current task, safety, efficie
SYSTEM STATE
TARGET STATE
what's the effect?
interpr. in terms of task
ident. in terms of deviation from the syst. state
which is the appropriate change in oper. cond.?
KNOWLEDGE-BASED ANALYSIS
KNOWLEDGE-BASED PLANNING
IDENTIFY present state of system
DEFINE TASK select approporiate change of syst. cond.
ident. in terms of task
SET OF OBSERV.
perceived as system state
ident in terms of proc. (e.g. via instructions)
TASK
what lies behind?
how to do it?
OBSERVE information and data
perceived in terms of task
FORMULATE PROCEDURE plan sequence of actions
what's going on?
preceived in terms of action (e.g. via prelearned cue)
ALERT
PROCE-DURE
interrupt in terms of time for task
ACTIVATION Detection of need for data processing
release of preset response
EXECUTE coordinate manipulations
RULE-BASED SHORTCUTS

Figure 6. A model of human decision sequence. (Rasmussen, 1981)

It is important to understand that without exception these methods rely on subjective estimation by experts for the data of which they make use. Several give the impression of considerable accuracy, since their outputs are almost all probabilities in the range (0,1), and measurement seems to be on a ratio scale. In every case the predictions can only be as good as the initial estimates. While it is to be hoped that empirical data on errors will become available at some time in the future, and can be used in these models, very few such data are presently available. Such data as are available are inadequate to support such estimates.

Figure 6. A Model of Human Decision Sequence. (Rasmussen 1981)

Throughout this section we shall refer to "models" of error. There is some doubt as to whether certain of the techniques involve modelling, but the term is convenient and little would be gained by arguing the point.

1) THERP: Technique for Human Error Rate Prediction.

THERP is the best known model. It was developed by Sandia Laboratories in the U.S. for use in their nuclear weapons manufacturing program, and although various people have been associated with its development, it is commonly identified as the work of Swain. (See e.g. Swain and Guttmann, 1982, NUREG/CR-1278). THERP makes use of an event-tree analysis of the tasks involved in nuclear plants, and its output is behaviourally oriented. The probabilities are probabilities ranging from incorrectly reading a meter, to failure to perform high level procedural actions such as trip the reactor. Performance Shaping Factors (PSFs) are used to modify the probabilities. These range from purely psychological factors such as fatigue, stress, set, expectation, time pressure, etc. to purely external features such as control room hardware (indicators, controls, and the like). The magnitude of the PSFs effect is in some cases very large, changing the probability of error by several orders of magnitude.

Perhaps because THERP has actually been used to perform human reliability analysis for PRAs in the nuclear industry it has been subject to particularly harsh criticism. Among the problems are the assumptions of independence when combining probabilities, omission of common-cause human errors in earlier versions, and the fact that the PSFs (and some of the probabilities) were chosen by a rather weak use of a Delphi like technique using a very small number of subject matter experts. On the credit side it must be said that the very small amount of empirical data collected by General Physics Corporation in recent years does seem to agree with the estimates of THERP. The estimates of both model and data are only accurate to about plus or minus an order of magnitude. THERP is, when all is said and done, probably the best of the techniques currently available. There is some indication that THERP is unreliable in the sense of test-retest reliability, and is very difficult to use in the hands of people other than those who developed it.

2) SLIM-MAUD: Success Likelihood Index Methodology - Multi Attribute Utility Decomposition.

This technique has been developed by Embrey and his co-workers (Embrey et al, 1984). In principle it can predict cognitive as well as behavioural errors. Its output is a probability of success (i.e. 1 - p(failure)) and an estimate of its range. It has been implemented on microcomputers as almost a "turn-key" system. MAUD is a computer program which conducts a dialogue with the user, leading the latter to discover the characteristics of people and tasks which the user believes to be most relevant to predicting error. This part of the technique is based on well understood and well established decision analysis techniques. Having clarified his ideas, the user then estimates the effects of the dimensions identified by MAUD as input to SLIM, which provides an estimate of the likelihood of success. Using the assumption that log(probability of error) is a linear function of Likelihood Index, estimates of error probability can be obtained by reference to two calibration tasks whose error probabilities are objectively known.

SLIM-MAUD is easy to use and is well accepted by users. At present only preliminary tests of its reliability and validity have been attempted. It suffers from the basic problem that it is using the estimates of subject matter experts as its input. It provides a somewhat rigorous and objective way to extract such estimates, but subjective estimates they remain. Preliminary data on validation were disappointing (Embrey et al, 1984), but the creators of SLIM-MAUD intend further development. If subjective estimates are an acceptable basis for predicting error, SLIM-MAUD has the potential to use them effectively to predict cognitive errors. The potential has yet to be realized in actuality.

3) OATS: Operator Action Trees

Developed by Wreathall at NUS Corporation in the U.S. this technique is based on an event tree approach. A generic example of the tree is shown in Figure 7. Unlike similar event trees already developed for PRA, the intent in OATS was specifically to include the operator's actions in the tree. Moreover as can be seen from the figure, cognitive actions such as diagnosis and choice are included. OATS suffers from the same basic problem as THERP in that the probabilities at present must be obtained from subject matter experts or extrapolated from unrelated data in the field, simulator or laboratory studies. OATS has not been formally validated, but it has been related to empirical data in a particularly interesting way by use of generic "time-to-completion" curves. As a result, OATS can predict the probability that no response will yet have been made to the annunciation of an incident as a function of time since the incident occurred. The resulting curves seem to provide a reasonable estimate of "speed accuracy trade-off" in fault diagnosis when compared to the data recently obtained by Beare et al (1983) and Kozinsky et al (1984). See Hall, Fragola and Wreathall (1982). See also the next section on "Time Reliability Curves".

Schurman and Banks (1985) suggest that OATS is particularly useful for cognitive and diagnostic behaviours rather than procedurally based tasks. Although OATS has not to date been validated it is a method whose development should be followed.

4) Time reliability curves.

Strictly speaking this method, though closely related to OATS is not a theory or model of error, but uses estimates of errors to predict the time to complete actions correctly. Beare et al. (1982, 1983), showed that there was a strong relationship between the time since a transient was annunciated and the probability that correct diagnosis and action were completed. Hannaman, Spurgin, and Lukic (1984) give a particularly clear account of the method. They propose that different kinds of cognitive activity can be completed at different rates. These rates are affected by different kinds of errors which may occur, including misperception, misdiagnosis, etc. An error will increase the time to completion. Using existing data (or estimates) of error probabilities and times to completion it is possible to relate probability that a task will be completed to the time

Figure 7.

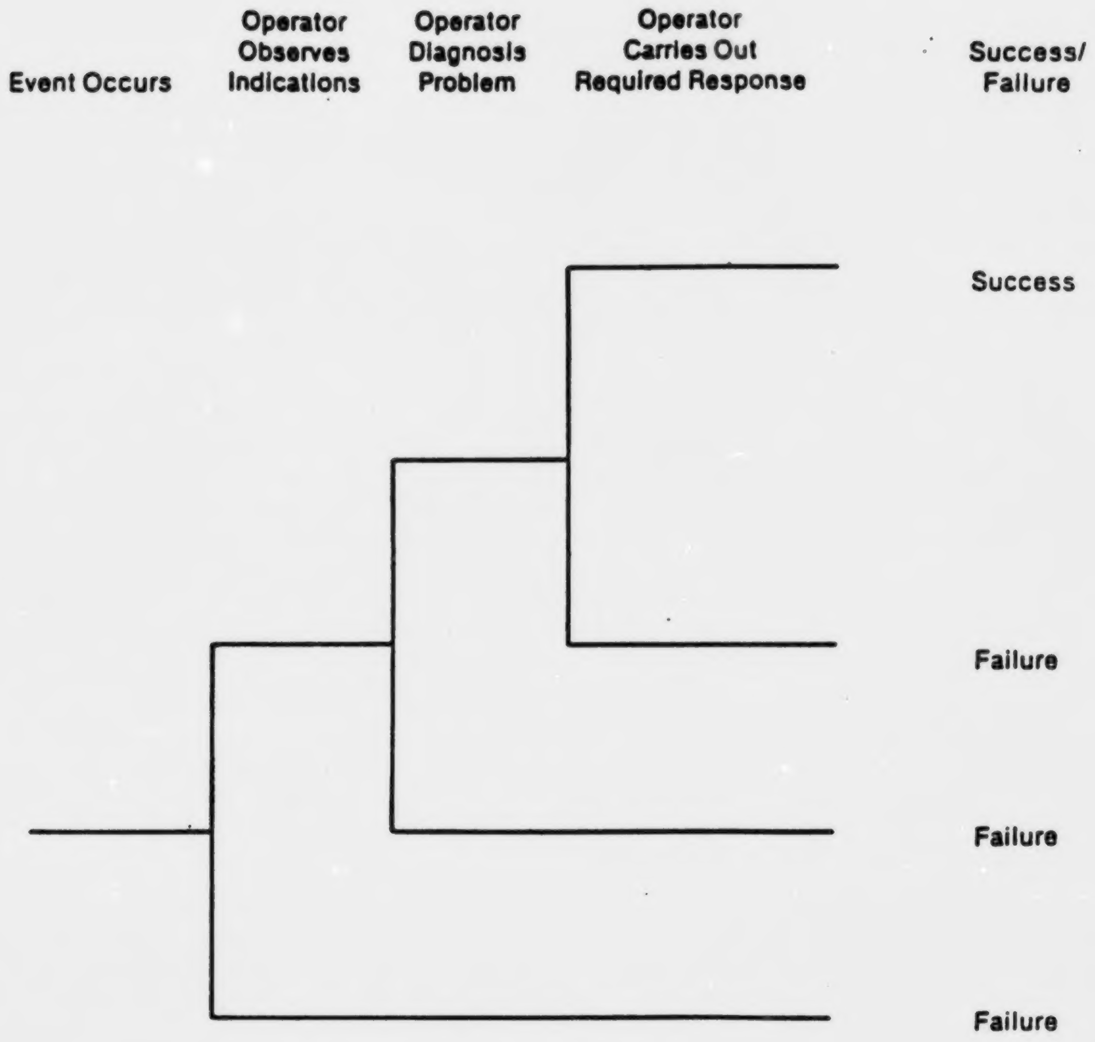

Event Occurs
Operator Observes Indications
Operator Diagnosis Problem
Operator Carries Out Required Response
Success/ Failure
Success
Failure
Failure
Failure

Figure 7. Example of the basic Operator Action Tree.

since annunciation using a three-parameter Weibull distribution. The parameter values depend on the type of task.

An excellent fit between the model and observed completion times can be obtained. Many such curves will be found in Hannaman, Spurgin and Lukic (1984). But while this model predicts time to task completion, and is a very important technique for that purpose, it is not a theory or model of error. It uses models or data on errors as its input, rather than producing such data as an output. Specifically, the occurence of an error increases time to completion: but the model does not predict when or how often such an error will occur. See also comments under section 3.2.4.

5) SHARP: Systematic Human Action Reliability Procedure

SHARP is actually a method for performing Human Reliability Analysis (HRA) using THERP and OAT models, and thus will still depend on subject matter expert judgement for input. SHARP provides a framework for the use of these models within a systematic and regularized PRA. The framework is a logic tree that explicitly identifies and screens human error as a significant risk factor. The framework is also provided by a step-by-step procedure for conducting the PRA. SHARP was developed by analyzing the steps actually taken in performing PRAs and deducing their common elements, then adding steps that should occur, whether or not they were common to the sample PRAs. An IEEE Guideline for HRA using SHARP is presently under development.

6) STAHR: Socio-Technical Assessment of Human Reliability.

This is another model in the early stages of development. It uses decision-analysis techniques and an influence diagram to incorporate some very complex variables into the prediction of error. For example the effects of morale, stress, teams vs. individuals, organization, etc. are added to the effects of the design of the control room, quality of EOPs, etc. It appears that subject matter expert consensus is the basis of the estimates used in the model. Schurman and Banks (1985) believe it will be of more value as applied to a specific NPP rather than as a generic predictor. Note that it was developed by Phillips, Humphreys and Embrey (1983) who also incorporated decision analysis techniques into SLIM-MAUD.

7) SAINT: (Systems Analysis of Integrated Networks of Tasks)

Siegel and Wolf (1967) developed a behaviouristic simulation of human operators alone or in teams of up to several people. The simulation technique, which is more properly regarded as a computer language than a model, is called SAINT and has been extensively used in military system development. The simulation is a network of activities which collectively describe a task. Each node is a particular element of the task. The simulator steps along a time-line performing the activity required by each node. The time taken to perform each node is described by a probability distribution, as is the accuracy with which the step is performed. Thus the simulation will produce error distributions, speed-accuracy trade-offs, distributions of time to completion, queuing or error when two things must be done simultaneously, etc. Because of the built in distributions the model can be run repeatedly as a Monte Carlo simulation to obtain distributions of outcomes, error, times, etc. See Siegel and Wolf (1967) and MAPPS (below).

Potentially these simulations could be applied to the modelling of cognitive errors. The simulation has been extensively validated in military contexts. But it was developed as a very behaviouristic model, and it is not clear where estimates of error and time distributions for cognitive tasks can be obtained. For behavioural tasks there are plausible candidate distributions in the psychological literature. Where the task is strongly determined by procedures, applications of Siegel-Wolf techniques might be relatively straight forward. There are rumors from time to time that cognitive factors will be included in SAINT, but at present it appears that they have not been so included in the sense which we believe is required be done to model cognitive error. The success of SAINT as a method for predicting cognitive error depends on first, an accurate cognitive task analysis, and second, accurate identification of the relevant probability distributions.

8) MAPPS: Maintenance Personnel Performance Simulation

MAPPS is a Siegel-Wolf SAINT model applied specifically to the behaviour of maintenance personnel. It is currently in the process of evaluation, but the document cited provides an example of how such simulation might be used to model behaviour of control room personnel. Note that the accuracy of the output depends on providing the simulation with appropriate data.

Both SAINT and MAPPS merely turn the crank of a well defined system. They will provide stochastic estimates of the interactions among events in a complex system, where the complexity is such that intuitive or manual calculations cannot be performed. The accuracy of the output is entirely dependent on the parameters of distribution supplied, and if they in turn come from subject matter experts, the user should beware of spurious appearances of accuracy and sophistication.

8) PROCRU: Supervisory Control Model

PROCRU is not a true acronym but the name of a computer program. It stands for Procedure Oriented Crew Model. This model has been used to simulate behaviour in a three person aircraft cockpit and has been proposed by its authors for use in modelling NPP Control Room activity (Baron et al (1982)). It is conceptually completely different from all the other models discussed above, and is a conceptually very sophisticated quasi-analytic model. Three kinds of behaviour are modelled, control, cognitive, and communication. The operators are modelled as if they assess the values of the system state variables using Optimal Estimation Method based on Kalman Filtering. The estimate so obtained can be used directly to choose appropriate (manual) control activities, using Optimal Control Theory. That sequence is most appropriate for modelling (perceptual-motor) skill-based behaviour. The same state variable information can, however, be used alternatively as input to cognitive decision mechanisms, as can the (verbal) inputs from the communications system. Verbal, qualitative and quantitative information is used by the model to generate decisions (such as a choice of procedure to use) and verbal messages to colleagues. These activities are modelled using an artificial intelligence approach using an Expert System technique. The Optimal Estimation and Control parts of the model can generate analytically error probabilities and probabilities of delays in response due to attention switching, that is, without the use of empirically determined or subjectively estimated values. The expert system portion of the model requires empirical data for its construction.

The output of the model when used to simulate aircrew is impressive. To apply it to NPPs would require great sophistication on the part of the user and large and sophisticated computer facilities. PROCRU or its extension cannot be regarded as a practical approach to error prediction at present, but it is in many aspects by far the most interesting model. It is the only one which uses a model of process at a high level to generate events at a low level. In all the other cases the data, model, and outputs are at similar levels. Personal communications from its authors indicate that development for NPP use is continuing.

3.4.5 Overview of models

It is clear that none of the above approaches satisfies the need for a predictive model of cognitive error in NPP operators. With the exception of PROCRU, which fails for its own reasons, all the models are little more than generalizations from data, and the data are in all cases of poor or very poor quality,and mainly subjective. There is nothing inherently wrong with using subjective judgements in measurement if it is done properly. (It should be noted that a well known document by Seaver and Stillwell (1983) describes a defective method that can give probabilities greater than 1.0: a revision published in 1985 appears to leave this deficiency uncorrected.) What is really required is an model which can generate estimates of error probabilities "top-down", analytically, independent of data. What is required is a model of process, not a model of data. Only with such a model, coupled with a detailed cognitive task analysis, will it be possible to predict events and distributions of events which occur so rarely that the collection of empirical data is impossible. No such method is to be found among those currently available.

Schurman and Banks (1985) provide a table for assessing the merits of the models of error outlined above. The dimensions used are cost, breadth of applicability, ease of use, face validity, precision, acceptance by professionals of quality of data, sensitivity to differences in systems being evaluated, predictive validity, and type of measurement. (The last category scores higher the more closely the method approaches a true ratio-scale measure.) In addition the method must be replicable (give the same results when used by different operators), and "public and scrutible". This last criterion requires that any user may see what the method involves. It must have no hidden assumptions, no special operations known only to those expert in its use, and be fully documented.

On such an assessment, Schurman and Banks find no method scoring above 50% of the possible score using a quantitative rating scale.

3.5 Taxonomies

3.5.1 The Nature of Taxonomies

Taxonomies classify behaviour without explicitly explaining it; an underlying model may or may not exist. While a model aims at explaining behaviour, a taxonomy classifies behaviours in a practical situation.

The taxonomies discussed below are all based on the information processing model that a human perceives, interprets and then acts. These divisions of behaviour are not completely separate -- for example in perceiving one acts by moving the eyes, in acting one perceives the position of the switch to be moved. This lack of separation is evident in the taxonomies described.

Taxonomies can be classed as behaviour oriented or task oriented or generic. Taxonomies of the first type include those described by Berliner et al (1964), Altman (1967) and Norman (1981). Those of the second type are typified by Rasmussen et al (1981). Senders (1983) describes a generic taxonomy.

In using a taxonomy to describe cognitive behaviour in a situation where errors are the main interest, two approaches may be taken. One can argue that distinct types of errors or error mechanisms exist, such as omission, (Senders 1983) or stereotype fixation, (Rasmussen et al., 1981), and the errors or error mechanisms themselves should be classified. The Rasmussen and Senders taxonomies are of this type (see Figures 8 and 9). Another approach is to argue that since error is very close to the desired correct behaviour, structurally, spatially, organically, etc., the behaviour leading to the error should be classified. The Berliner taxonomy is of this type (see Figure 10).

3.5.2 Behaviour Oriented Taxonomies

The Berliner taxonomy (Berliner et al, 1964) describes behavioural mechanisms which produce error if they do not function properly. The taxonomy is structured in a hierarchical fashion, the first level describing processes (perceptual, cognitive, motor and communication), the second level describing activities associated with the processes, (e.g. the identification of objects, actions and events is an activity associated with the perceptual process), the third level describing specific behaviours that occur during the activities, (e.g. interpolating, verifying or remembering during the activity of information processing). The fourth level gives definitions for each of the specific behaviours.

Though perceptual, motor, and communication processes all have cognitive elements, the cognitive category refers to purely cognitive behaviours which occur before or after the other processes. The Berliner taxonomy is oriented towards describing human behaviour involved in carrying out tasks; it is not specifically directed towards identifying the nature of human error while carrying out a task.

Reason (1977) developed a taxonomy of behaviour leading to error by having 35 subjects keep diaries of absent minded errors for a two week period. This generated 433 incidents. The errors were then classified using the algorithm referred to earlier and shown in Figure 4. By examining where errors occurred, for example in the formation of the intention, in the feedback from the action system to the intention system, in the selection of the action sequence etc., Reason was able to classify 95% of the errors. Errors relating to forgetting the plan or the preceeding actions were classified as storage failures. These were found to be the most common class of errors (41.4%). Selection failures, where an inappropriate action sequence was substituted for the intended one, made up the second largest class of errors -- 37.2%. Table 1 shows the categorization of the various errors.

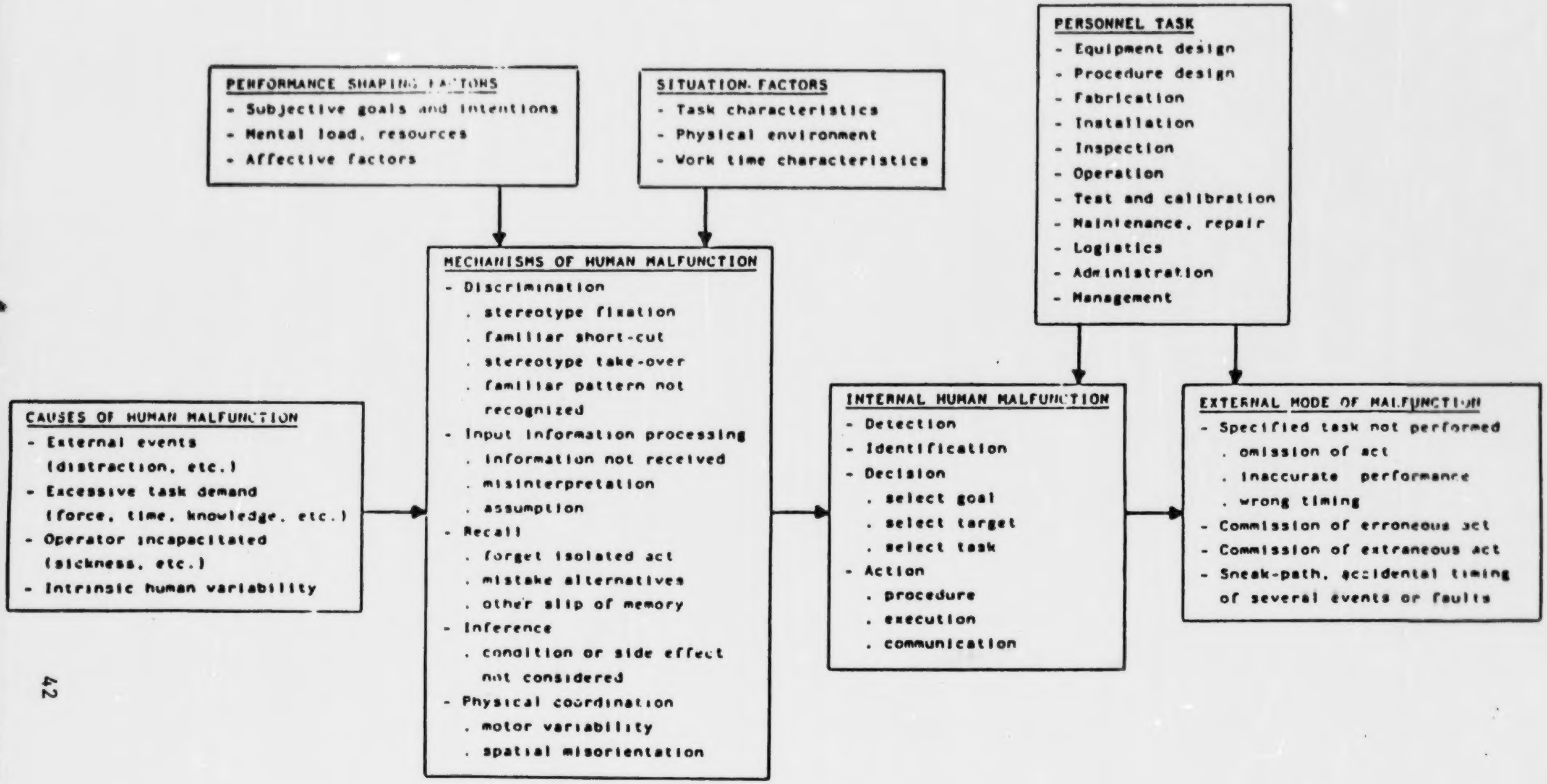

PERFORMANCE SHAPING FACTORS
- Subjective goals and intentions
- Mental load, resources
- Affective factors
SITUATION FACTORS
- Task characteristics
- Physical environment
- Work time characteristics
PERSONNEL TASK
- Equipment design
- Procedure design
- Fabrication
- Installation
- Inspection
- Operation
- Test and calibration
- Maintenance, repair
- Logistics
- Administration
- Management
CAUSES OF HUMAN MALFUNCTION
- External events (distraction, etc.)
- Excessive task demand (force, time, knowledge, etc.)
- Operator incapacitated (sickness, etc.)
- Intrinsic human variability
MECHANISMS OF HUMAN MALFUNCTION
- Discrimination
. stereotype fixation
. familiar short-cut
. stereotype take-over
. familiar pattern not recognized
- Input information processing
. information not received
. misinterpretation
. assumption
- Recall
. forget isolated act
. mistake alternatives
. other slip of memory
- Inference
. condition or side effect not considered
- Physical coordination
. motor variability
. spatial misorientation
INTERNAL HUMAN MALFUNCTION
- Detection
- Identification
- Decision
. select goal
. select target
. select task
- Action
. procedure
. execution
. communication
EXTERNAL MODE OF MALFUNCTION
- Specified task not performed
. omission of act
. inaccurate performance
. wrong timing
- Commission of erroneous act
- Commission of extraneous act
- Sneak-path, accidental timing of several events or faults

Figure 8. Multifacet taxonomy for description and analysis of events. (Rasmussen, 1981)

		Locus of Human Error		Type of Error			
		Endogenous	Exogenous	Omission	Insertion	Substitution	Repetition
PROCESS	Perception						
	Intention						
	Execution						
SOLUTION		Training Selection Redundancy	Human Factors Engineering				

Figure 9. Senders' classification scheme for Error.

PROCESSES	ACTIVITIES	SPECIFIC BEHAVIORS	DEFINITIONS
1. Perceptual	1.1 Searching for and Receiving Information	1.1.1 Inspects	To examine carefully, or to view closely with critical appraisal.
		1.1.2 Observes	To attend visually to the presence or current status of an object, indication, or event.
		1.1.3 Reads	To examine visually information that is presented symbolically.
		1.1.4 Monitors	To keep track of overtime.
		1.1.5 Scans	To quickly examine displays or other information sources to obtain a general impression.
		1.1.6 Detects	To become aware of the presence or absence of a physical stimulus.
	1.2 Identifying Objects, Actions, Events	1.2.1 Identifies	To recognize the nature of an object or indication according to implicit or predetermined characteristics.
		1.2.2 Locates	To seek out and determine the site or place of an object.
2. Cognitive	2.1 Information Processing	2.1.1 Interpolates	To determine or estimate intermediate values from two given values.
		2.1.2 Verifies	To confirm.
		2.1.3 Remembers	To retain information (short-term memory) or to recall information (long-term memory) for consideration.
	2.2 Problem Solving and Decision Making	2.2.1 Calculates	To determine by mathematical processes.
		2.2.2 Chooses	To select after consideration of alternatives.
		2.2.3 Compares	To examine the characteristics or qualities of two or more objects or concepts for the purpose of discovering similarities or differences.
		2.2.4 Plans	To devise or formulate a program of future or contingent activity.
		2.2.5 Decides	To come to a conclusion based on available information.
		2.2.6 Diagnoses	To recognize or determine the nature or cause of a condition by consideration of signs or symptoms or by the execution of appropriate tests.
3. Motor	3.1 Simple/Discrete	3.1.1 Moves	To change the location of an object.
		3.1.2 Holds	To apply continuous pressure to a control.
		3.1.3 Pushes/Pulls	To exert force away from/toward the actor's body.
		3.1.4 Attaches	To affix an object to a larger object by tying or gluing.
		3.1.5 Gives	To put an object into the possession of another for his use.
		3.1.6 Removes	To detach and move out of position.
		3.1.7 Discards	To get rid of an unnecessary or useless object.
	3.2 Complex/Continuous	3.2.1 Positions	To operate a control that has discrete states.
		3.2.2 Adjusts	To operate a continuous control
		3.2.3 Types	To operate a keyboard.
		3.2.4 Installs	To put into an appointed place or position.
4. Communication		4.0.1 Answers	To respond to a request for information.
		4.0.2 Informs	To impart information.
		4.0.3 Requests	To ask for information.
		4.0.4 Records	To document something, as in writing.
		4.0.5 Directs	To ask for action.
		4.0.6 Receives	To be given written or verbal information
		4.0.7 Returns	To give an object back to its owner.

Figure 10. Berliner taxonomy.

Table 1

Presumed type of failure	Category	Examples and percentages
Storage	A	*Undetectable errors*—both the original intention and the failure to execute it are forgotten. (0)
Storage	B	*Forgetting items in the plan*—part of the plan is temporarily lost but later recalled. Example: forgetting to post a letter during a shopping expedition. (22.4)
Storage	C	*Forgetting the plan*—where the existence of the plan is not forgotten, indeed its actions are usually under way, but its substance cannot be recalled. Example: 'Went over to open the drawer, but forgot what I wanted'. (6.5)
Storage	D	*Forgetting preceding actions*—preceding actions are either forgotten or incorrectly recalled. Often results in losing one's place in the plan. Example: boiling a second kettle of water after the tea has already been made. Another common variant is forgetting where articles were put down. (12.5)
Test	E	*Verifications errors*—where a planned sequence is terminated too early or is continued beyond its intended conclusion. Example: getting into bath with clothes still on. (6.0)
Discrimination	F	*Classification errors*—input wrongly identified, often leading to actions appropriate for erroneous classification but not for ongoing plan. Example: taking out one's own front-door key on approaching a friend's house. (10.9)
Selection	G	*Branching errors*—where two different outcomes have initial actions in common, but the actions proceed towards the unintended outcome. Example on passing through back porch on the way to his car, one subject put on the clothes and wellington boots that are kept there for gardening (14.3)
Selection	H	*Misordering errors*—the correct actions are executed, but in the wrong order Example: while filling a bucket with water, subject replaced the lid before turning off the tap (6.0)
Selection	I	*Insertions and omissions*—when unwanted actions are added or necessary actions omitted Example turning on a light when leaving a room in daylight; filling electric kettle switching on socket, but failing to insert plug (12.5)
Selection	J	*Corrected errors*—where a sequence of actions initially deviates from plan, but is later corrected. Example: starting to look for a jar of coffee in a cupboard and then realizing it is kept in pantry (1.2)
Selection	K	*Total errors*—where all actions are inappropriate for plan Example trying to light an electric fire with a box of matches (3.2)

Table 1. Examples and relative frequencies of the categories of actions not-as-planned. (From Reason, 1976)

Norman (1981) classifies errors as "mistakes" or "slips". "Mistakes" are errors of intention, whereas "slips" are errors in carrying out the intention, i.e. errors in action. Norman describes a behaviour-oriented taxonomy which was developed after reviewing 1000 instances of 'action slips' or errors that occur when one carries out an unintentional action. Norman considers actions to be sequences of activation-trigger schemata. Three major sources of action slips are classified as formation of the intention, triggering and activation. Table 2 shows a classification of slips based on heir presumed sources.

Slips incurred during formation of an intention include errors in classifying the situation and errors resulting from ambiguous or incompletely specified intentions. The activation of a schema can be faulty in that it is unintentionally activated or that it loses its activation before it can be run off, so that an omission occurs.

Unintentional activations can result from capture slips -- an old habit takes over inappropriately, or from 'data driven' slips where interaction with the external environment causes the slip. Slips resulting from faulty triggering include anticipation errors where thoughts are confused with deeds and an act is omitted.

A feedback mechanism is postulated which allows people to detect their errors. The mechanism includes feedback at different levels to monitor the many processes involved in carrying out an intention.

Altman (1967) developed a behaviour-oriented taxonomy. Like the Berliner, this taxonomy is also set up in a hierarchical fashion. The first level (column) describes behaviours as they increase in complexity, starting with A) sensing, detecting, etc, B) rote sequencing, C) estimation with discrete or continuous responding, D) rule using and decision making, E) problem solving. At each level modes of error possible are specified (second column). For example in rote sequencing errors of omitting a procedural step, making a procedural step in the wrong order and so on are listed. Types of learning or training involved at each level of behavioural complexity are shown in a third column. Moray (1981) added a fourth column in which relevant psychological research is identified. (See Figure 11.)

Moray also emphasizes that as the level of behavioural complexity increases, from sensing, detecting, etc. through to problem solving, there is increasing probability that individual differences and experience play a significant role. The scope and cost of errors also increase as does the difficulty of recovering from the error.

Another scheme for classifying human behaviour leading to error was developed by Senders (1983) (See Figure 9). The scheme classifies errors by locus, by process, by expression and by remedy. The locus of error is endogenous, that is, within the human, or exogenous, outside the human. If an operator incorrectly determines the system state on the basis of insufficient information then the error would be endogenous. If a communication system is noisy, with a poor signal to noise ratio, even an optimal listener will make errors. These would be exogenous. In general, exogenous errors call for system redesign, endogenous errors for retraining. This gives rise to the classification of error according to how to deal with it. In the endogenous case training, selection and luck are required, in the exogenous case, human factors engineering is called for.

TACTICAL ← INFORMATION PROCESSING → STRATEGIC

INCREASING PROBABILITY THAT INDIVIDUAL DIFFERENCES AND EXPERIENCE PLAY A SIGNIFICANT ROLL →

INCREASING SCOPE OF EFFECTS AND COSTS OF EFFECTS ASSOCIATED WITH ERROR AND INCREASING DIFFICULTY IN RECOVERING FROM ERROR

CLASSIFICATION OF ERRORS:

Level of Behavioural Complexity	Modes of Error	Types of Learning or training involved	Relevant Psychological Data from other research needed to understand origin of error
A. Sensing, detecting, identifying, classifying.	1. Failure to monitor the environment. 2. Failure to report or record a change. 3. Non-existent change reported or recorded. 4. Failure to detect a target. 5. Non-existent target detected (false alarm). 6. Assignment of observed change or target to the wrong category or class.	Perceptual Learning Probability Learning Payoff Matrix Learning Concept Formation	Arousal Theory, fatigue,sampling behaviour, signal detection theory, speed-accuracy trade-off, man as an intuitive statistician, models of data acquisition, sequential decision theory, attention theory, monitoring and vigilance theory.
B. Chaining or rote sequencing.	1. Making a substandard response. 2. Omitting a procedural step. 3. Inserting an unnecessary procedural step. 4. Making a procedural step in the wrong order.	Rote Learning. Associative Chaining Mnemonics Cognitive Mazes	Set and Attention, Memory span, practice curves and extinction of learning, long term memory access and retrieval.
C. Estimation with discrete or continuous responding (Button/Key Press or tracking).	1. Failure to respond to above threshold change. 2. Response to below threshold change. 3. Too early response to change. 4. Late response to change. 5. Wrong direction of response. 6. Magnitude of response too great. 7. Magnitude of response too small. 8. Response continued too long. 9. Response not continued long enough. 10. Incorrect choice of transfer function. 11. Bandwidth too high. 12. Bandwidth too low.	Motor Learning. Perceptual-Motor Skills. Running Memory Span. Extrapolation from Data. Internal Models of System Dynamics. Generalistion.	Arousal, fatigue, man as statistician, running memory span, signal detection theory,reaction times, open and closed loop skills, sequential decision making, classical and optimal control theory, optimal estimation theory, Human bandwidth Human Transfer Function, speed-accuracy tradeoff,thresholding information theory attention, probability learning,adaptive control theory, motor programs monitoring and vigilance theory.

Figure 11. Altman/Moray scheme for classification of error.

TACTICAL — INFORMATION PROCESSING — STRATEGIC

INCREASING PROBABILITY THAT INDIVIDUAL DIFFERENCES AND EXPERIENCE PLAY A SIGNIFICANT ROLE

INCREASING SCOPE OF EFFECTS AND COSTS OF EFFECTS ASSOCIATED WITH ERROR AND INCREASING DIFFICULTY IN RECOVERING FROM ERROR.

2.

Level of Behavioural Complexity	Modes of Error	Types of Learning or training involved	Relevant Psychological Data from other research needed to understand origin of error
1. Logical manipulation, Rule using, & decision making.	1. Incorrect weight given to componunts of response contingencies. 2. Failure to apply an available rule. 3. Application of correct but inappropriate rule. 4. Application of incorrect rule. 5. Failure to obtain or use all relevant information. 6. Failure to identify all reasonable alternatives. 7. Making an unnecessary or premature decision. 8. Delaying a decision beyond time when needed. 9. Assuming incorrect possession of adequate knowledge. 10. Assuming incorrect possession of adequate skill.	Rule Learning Learning to Learn Reasoning Heuristics and Algorithms Unlearnt Behaviour Ill Defined Problems Innovation and Divergent Thinking.	Payoff estimation, probability estimation, speed accuracy trade-off, internal models, man as statistician, set, cognitive tunnel vision, working memory and long-term memory access, logkeeping, arousal and fatigue. Decision theory, subjective values, social variables, fuzzy set models and production systems.
2. Problem Solving.	1. Formulate erroneous rules or heuristics. 2. Failure to use available information to derive solution. 3. Acceptance of inadequate or suboptimal solution as final. 4. Incorrect Assumption of adequate knowledge. 5. Incorrect assumption of adequate skill.		

Figure 11. Altman/Moray scheme for classification of error.

Table 2

- Slips that result from errors in the formation of the intention
  - Errors that are not classified as slips: errors in the determination of goals, in decision making and problem solving, and other related aspects of the determination of an intention
  - Mode errors: erroneous classification of the situation
  - Description errors: ambiguous or incomplete specification of the intention
- Slips that result from faulty activation of schemas
  - Unintentional activation: when schemas not part of a current action sequence become activated for extraneous reasons, then become triggered and lead to slips
    - Capture errors: when a sequence being performed is similar to another more frequent or better learned sequence, the latter may capture control
    - Data-driven activation: external events cause activation of schemas
    - Associative activation: currently active schemas activate others with which they are associated
  - Loss of activation: when schemas that have been activated lose activation, thereby losing effectiveness to control behavior
    - Forgetting an intention (but continuing with the action sequence)
    - Misordering the components of an action sequence
    - Skipping steps in an action sequence
    - Repeating steps in an action sequence
- Slips that result from faulty triggering of active schemas
  - False triggering: a properly activated schema is triggered at an inappropriate time
    - Spoonerisms: reversal of event components
    - Blends: combinations of components from two competing schemas
    - Thoughts leading to actions: triggering of schemas meant only to be thought, not to govern action
    - Premature triggering
  - Failure to trigger: when an active schema never gets invoked because
    - The action was preempted by competing schemas
    - There was insufficient activation, either as a result of forgetting or because the initial level was too low
    - There was a failure of the trigger condition to match, either because the triggering conditions were badly specified or the match between occurring conditions and the required conditions was never sufficiently close

Table 2. Categorization of action slips.
(From Norman, 1980)

The third level of classification is the fine structure of the error. Errors may occur in any of three processes: perception, intention, and execution. The final level of this scheme describes the way in which an error is expressed. The four possibilities are omission, substitution, insertion and repetition. Omission of a step may result from an error in perception or intention. Substitution may occur due to errors in perception, intention or execution, while repetition is an error in perception or execution but not intention. Note that if there is an error in more than one step, correct behaviour may result because the second error by chance counteracts the first. The probability of such accidental corrections will be extremely low, and in general the second order probabilities may be ignored.

3.5.3 Task Oriented Taxonomies

Rasmussen (1981) describes a task oriented taxonomy which is specifically directed toward the description of situation characteristics and analysis of events involving human error. The taxonomy was developed using categories of error found typical from knowledge of human information processing and performance shaping factors.

Various categories to describe the situation and the events are given. Situation characteristics considered are performance shaping factors such as subjective goals and intentions, situation factors such as work-time characteristics, and the personnel task, for example test and calibration. Categories related to analysis of the event leading to human error describe the causes of human malfunction (for example excessive time demand), the mechanisms of human malfunction, (for example in the detection process), and the external mode of malfunction, (for example the commission of an erroneous act). The taxonomy is shown in the form of a block diagram with the performance shaping and situation factors affecting the mechanism of human malfunction, and the personnel task affecting both the internal malfunction and the external mode of malfunction. (See Figure 8).

The Rasmussen taxonomy does not separate purely cognitive from partially cognitive (perceptual, motor and communication) processes as does the Berliner, with the result that the "internal human malfunction" category contains some perceptual, some cognitive and no communication or motor behaviours. For example the activities of inspecting, calculating, informing, etc. are not included. Some of the behaviours not included are important in error producing situations. For example, there is documented evidence of operator errors arising during emergency procedures due to poor mental arithmetic, i.e. calculating, and poor communication. Neither of these behaviours is included in Rasmussen's taxonomy.

Activities such as decision making are treated differently in each taxonomy. The Berliner taxonomy distinguishes between different levels of decision making: choosing ('to select after consideration of alternatives'), deciding ('coming to a conclusion based on available information'), and diagnosis ('recognizing or determining the nature or cause of a condition by consideration of signs or symptoms or by the execution of appropriate tasks'). In contrast, Rasmussen refers to decision in terms of selecting a malfunction" are not connected, though this might help in the analysis of events. For example the activity of detection might result in error as a consequence of inattention or stereotype takeover.

While Rasmussen's taxonomy has had an enormous heuristic influence on research in nuclear and other process control industries, it does not provide a way to predict error quantitatively. It stimulates conceptual analysis and it guides research. It does not solve the problem facing us here.

Beattie and Iwasa-Madge (1984) describe a taxonomy which was developed to facilitate human reliability analysis in a probabilistic safety evaluation of a nuclear power plant, by including human-machine interactions in fault trees. Only human-machine interactions were considered, thus communication errors are excluded. Interactions were divided into simple and complex according to the ease with which they could be assigned error rates. Simple interactions were classified as being based on written or learned procedures, involving interaction with a control mechanism (switch etc.) or display, and occurring prior to an initiating event. The remaining interactions were considered to be complex. Complex interactions were further subdivided into those which occured prior to an initiating event, where behaviour is rule based, and those which occured after an initiating event where knowledge based behaviour predominates.

All three types of interactions were divided into two categories; the first being disability interactions, which include errors where the operator fails to return a workable component/system to the state required to perform its intended function, and the second failure detections, where the operator has not detected a failed component/system. The taxonomy is shown in Table 3.

For the simple human interactions and the pre-initiating event complex interactions the taxonomy is set up to describe the phase of operation during which the error occurred rather than to describe the reason for its occurrence. The authors state that the simple human interactions require only selection of a probability and a few modifying factors, presumably by applying a fault tree model to each kind of simple interaction.

3.6 Cognitive Engineering

One of the long term goals of the study of cognitive behaviour is clearly better design of the human-machine interface, through "cognitive engineering", i.e. engineering which takes into account the complex motivations involved in cognitive behaviour. Hollnagel and Woods (1983) discuss the limitations of the current behaviouristic understanding of humans in design of human-machine systems where supervisory rather than manual control dominates. According to these authors "man functions according to a psychologic rather than to a logic". Thus psychological theories of decision-making such as that developed by Janis and Mann (1977) provide more accurate descriptions of behaviour than do descriptions of humans as mathematically rational decision makers, (e.g. the Bayesian statistics model.) Similarly with perception, psycho-physics rather than physics provides a more relevant description of humans. The authors discuss changes in the cognitive task of an operator whose plant is automated and how these changes are often poorly understood by the system designer. For example the introduction of VDTs has caused problems far more because of changes in job content than because of poor design per se. Hollnagel and Woods agree that the application of traditional human engineering in such systems may reduce particular errors but others may occur because the total effect of changes on the underlying cognitive system has not been considered.

The authors conclude that there is a need to develop methods for cognitive task analysis to identify the operator's internal model in order to provide the tools for the design of complex systems, and to allow cognitive engineering to be carried out.

FIGURE 2
THE TAXONOMY CONSISTS OF 3 TABLES:

	Pre-Initiating Event	Post-Initiating Event
"Simple" Interactions	Table 1	
"Complex" Interactions	Table 2	Table 3

TABLE 1
SIMPLE HUMAN INTERACTIONS

DISABILITY INTERACTIONS

I(a) Manually Operated Valve Left in Incorrect Position Following	(i) Mtce/Calibration During Shutdown, and Error Not Detected (ii) Mtce/Calibration During Routine Operations, and Error Not Detected (iii) Routine Testing, and Error Not Detected (iv) A Normal Operation, and Error Not Detected
I(b) Handswitch/Pushbutton Left in Incorrect Position Following	(i) Mtce/Calibration During Shutdown, and Error Not Detected (ii) Mtce/Calibration During Routine Operations, and Error Not Detected (iii) Routine Testing, and Error Not Detected (iv) A Normal Operation, and Error Not Detected
I(c) Circuit Breaker/Wire Pair/ Connector Left in Incorrect Position Following	(i) Mtce/Calibration During Shutdown, and Error Not Detected (ii) Mtce/Calibration During Routine Operations, and Error Not Detected

FAILURE DETECTION INTERACTIONS

II(a) Components/System Failure Not Detected	(i) By Direct Indication from Display Devices in Main Control Room during Routine Operations

TABLE 2
COMPLEX INTERACTIONS – PRE-INITIATING EVENT

DISABILITY INTERACTIONS

I(a) System/Subsystem Left in Incorrect State Following:	(i) Mtce/Calibration During Shutdown, and Not Detected (ii) Mtce/Calibration During Routine Operations, and Not Detected (iii) Routine Testing, and Not Detected (iv) Normal Operation, and Not Detected

FAILURE DETECTION INTERACTIONS

II(a) No Response to Failed Component/System:	(i) By Interpretation of Display Devices in Main Control Room during Routine Operations

TABLE 3
COMPLEX INTERACTIONS – POST-INITIATING EVENT

DISABILITY INTERACTIONS

I(a) Omission – Required Action Not Taken, and Not Detected

I(b) Commission – Required Action is Taken Incorrectly, and Not Detected:
- Too Soon, Too Late
- Out of Sequence
- Too Little, Too Much

I(c) Commission – Action Taken When Not Required, and Not Detected
- Inhibition – Component/System Prevented from Functioning, in Error
- Inadvertency – Component/System Caused to Function, in Error

FAILURE DETECTION INTERACTIONS

II(a) No Response to Failed Component/System:	(i) By Display Devices in Main Control Room

Table 3. Taxonomy of NPP operator error.
(From Beattie and Iwasa-Madge, 1984)

Rasmussen (1983a) accepts the inevitability of some human error, even with the best human factors engineering. He suggests that errors are only errors because they occur in an environment which does not give sufficient feedback and is unforgiving. This assertion does not conform to the usage of virtually all other workers in the field. (Most workers in the field would argue that errors would remain errors irrespective of the environment.) It should be possible to correct the effects of inappropriate variations in performance before they lead to unacceptable consequences. Rasmussen says that reversability of error is largely a question of system properties while observability of error should be considered as part of the system design.

Norman (1983a) like Rasmussen discusses the need for "cognitive engineering" to help reduce the incidence of error and minimize its effect. He looks at computer system design and shows that the analysis of errors leads to several principles of design. These include the need for feedback, the need for dissimilar command sequences to carry out different classes of actions (to avoid "capture" or substitution errors), the need for reversability of actions (where possible) and the need for consistency in the structure and design of the system.

While the papers discussed above focus more on changes in machine design to take account of cognitive behaviour of operators, Sage (1981) examines changes in the human side of the system through training. Sage discusses numerous ways in which the human is not an optimal information processor or decision maker. To combat these deficiencies, which are discussed in detail later in this report, he suggests a number of means by which cognitive bias can be reduced and recognized. Those particularly applicable to training include:

1. Sample information from a broad data base and be especially careful to include data bases which might contain disconfirming information.

2. Encourage use of models and quantitative aids to improve information analysis through proper aggregation of acquired information.

3. Encourage decision makers to distinguish good and bad decisions from good and bad outcomes in order to avoid various forms of selective perception such as, for example, the illusion of control.

4. Encourage effective learning from experience. Encourage understanding of the decision situation and methods and rules used in practice to process information and make decisions such as to avoid outcome irrelevant learning systems.

5. Both qualitative and quantitative data should be collected, and all data should be regarded with "appropriate" emphasis. None of the data should be overweighted or underweighted in accordance with personal views, beliefs, or values only.

6. People should be reminded, from time to time, of the type or size of sample from which data are being gathered, so as to avoid the representativeness bias, a tendency to think that small samples are more typical of the properties of the whole than is in fact the case.

7. Information should be presented in several sequences so as to avoid the tendency to remember best those items which occur at the beginning and end of sequences.

3.7 The Assessment of Behaviour Involving Cognitive Error

3.7.1 Methods of Assessment

The methods for assessing cognitive behaviour of plant operators are discussed below with attention paid to the advantages and drawbacks of each method. First field analyses of data from significant event reports will be reviewed followed by nuclear power plant simulator studies, other complex plant simulator studies, and laboratory studies. Verbal report assessment techniques which are used in field, simulator and laboratory studies will then be discussed.

3.7.2 Significant Event Reports

Speaker et al (1983) describe an analysis of human error data contained in 8749 LER's, 3565 of which were from BWR's and 5184 from PWR's. Only 8% of the Licensee Event Reports (LER's) indicated human error as a cause. This is much lower than the usual figure given for operator error in complex systems and, given the degree of manual control in American BWR's and PWR's, is not credible. The authors note that there is a tendency to place the fault on a mechanical device or procedure even in cases where it was clear human error contributed to the event. A major problem cited with LER's is that they are neither trustworthy nor complete enough to permit identification of and analysis of human diagnostic error. Bearing these problems in mind the findings of the LER analysis were as follows:

1. Maintenance is at least as potent a source of unreliability as operations, and probably greater.

2. Omission errors account for about 50% of all errors. Since ommissions usually arise either from a failure to observe displayed information, failure to interpret that information, or failure to carry out the correct procedure, this figure emphasises the great importance of the design of procedures and manuals and of training both in their use, and in optimal monitoring behaviour.

3. Operator/maintainer tasks associated with testing (e.g. safety system testing) produced 6 to 7 times as many errors as tasks associated with normal operating.

3.7.3 Nuclear Power Plant Simulation Studies

Lees and Sayers (1976) describe experimental work on operator response to emergencies using training simulators for British nuclear power plants. In this work the fault condition, the allowable response time (90 seconds) and the corrective actions required were relatively well defined. For this reason the authors regard their estimates as an upper limit of operator reliability. Fault conditions simulated were control rod runout, coolant flow fault, blower failure, and gas temperature rise. Ten fault conditions were presented per hour, a high rate compared to the demand rate of about 0.0001 per hour for an industrial manual trip. Mean probability of failure was found to be 0.18. Operator reliability to complete the corrective actions increased with allowable response time, with high reliability being found for time greater than 60 seconds.

Because of the demand on nuclear power plant simulators as training devices, and the lack of automated data recording in many of them, only a few studies of operator performance in simulators have been conducted.

Crowe et al (1984) describe experiments carried out at Oak Ridge using the performance measurement system which is built into the PWR training simulator. The performance measures were supplemented by questionnaires which asked the operaters to rate tasks for difficulty. Fourteen events, from Large Break Loss Of Coolant Accident (LOCA) to expected normal transients were investigated. The crews were not station control room crews, but were ad hoc crews who happened to be training together. (Response time may be different with well integrated crews). The times measured are "the interval between the appearance of the first cue to the operators that something has happened and the first correct action they make in response to the malfunction". Times were given for crews as a whole, not individuals. The data did not include repeated performances by a single crew to the same kind of event: each crew experienced one of each kind of event.

There was more variation among teams than among events. Error rates and response times were not significantly correlated. Also response time was found to be unrelated to the operator's age, years of college education or power plant experience. However, teams with more experienced members made fewer omission errors.

Simulations of any kind have a number of drawbacks when used to study error. First, the simulation is never a completely accurate representation of the real process. In particular, time constants of the processes represented may be different. Second, some of the factors which distract the operator and contribute to error in real operating situations are carefully controlled out of simulations. These include interaction with others including operators, supervisors and maintenance personnel. Third, in order to collect any data the rate of presentation of emergency events must be several orders of magnitude greater than the rate that occurs in reality. The operator goes into the simulation expecting to deal with emergencies and can be expected to respond to the same with much more alacrity than usual.

As Van Gelder (1980) points out, the real value of simulation experiments can only be estimated in a satisfactory way when a broader field of real-life data, from critical incident reports for example, are available. Beare et al (1983) did do such a study, but looked at differences in response time between simulator experiments and Licensee Event Reports rather than error rates. In the simulator there was no difference in the times to initiate responses to step and to ramp failures, but the response time to ramp failures in the field were much longer than step failures. In general, for BWR data, response times for field events were nearly one log unit (10 times) as long as the simulator events. Although real response times may be very much longer than those found in simulators, a properly designed simulation can still contribute much to the prediction of human error in industrial processes.

3.7.4 Simulations of Plants Other Than NPPs

Researchers have also examined cognitive behaviour of operators controlling plants which require some of the same cognitive skills as nuclear power plants. Van Eekhout and Rouse (1981) examined seven crews of marine engineering officers who performed the task of coping with six different failures in a high-fidelity supertanker engine control room simulator. Measurement methods included verbal protocols, computer logs of all discrete events, interviews, questionnaires and observer ratings. Rasmussen's error classification scheme was modified by expanding the two or three categories that errors generally fell into. The categories used followed a natural flow from onset of the symptoms of the failure to compensate for the consequences. The relative roles of operator knowledge of the basic system and the properties of the automatic system, human factors design inadequacies and simulator fidelity inadequacies, were examined for each error.

It was found that errors associated with inappropriate identification of the failure were highly correlated with a lack of knowledge of the functioning of the basic system and the automatic controllers within the system. Further, errors related to execution of procedures were highly correlated with inadequacies of the layout of the control panel and simulator fidelity inadequacies. Based on these results, it was concluded that operator training should place increased emphasis on the knowledge necessary for dealing with failure situations.

Van Gelder (1980) used 12 apprentice operators of chemical process plants to study the effects of training and comprehensiveness of the task instruction on control of a simulated chemical process. A blending process was simulated on a CRT screen. An overall performance score for controlling the process was calculated by combining such measures as time to complete the task, sum of absolute deviations, number of time alarms were initiated, etc. Error rate was also recorded. While operators with the brief instructions had poorer performance scores and higher error rates for the first few tasks in the series compared to operators given comprehensive instructions, these differences had disappeared after practice. In a second series of experiments where process disturbances were introduced, probability of operator error varied between less than 0.001 for allowing vessels to run empty and for explosions, to 0.013 for interpretation.

Johnson and Rouse (1982) looked at the effects of different training methods on error rates in troubleshooting actual aircraft power plants. Subjects for the experiments were aircraft maintenance trainees. In the first experiment three training methods were used, one was context-free (TASK), the second context-specific (FAULT) and the third a video presentation. Performance was measured by having an evaluator complete a standardized check list. Data from the check-list were then used to calculate a performance index involving both quality of performance and time to complete diagnosis. Errors in diagnosis were classified using a modification of the scheme developed by Van Eekhout and Rouse (1981). (See Table 4.)

The FAULT training method was found to produce more failures of diagnosis. An analysis of the errors led to the conclusion that the subjects trained with FAULT knew what tests to make but did not necessarily know how to make them. This problem was corrected and a subsequent experiment showed that improving the training method had reduced the specific types of errors previously found. The authors emphasize that error classification schemes, to be useful, must be designed for the particular type of system being studied, or must be general, with numerous fine-grained levels, many of which will only be useful for particular types of tasks.

3.7.5 Laboratory Studies

Prior to the simulator experiments described earlier, Lees and Sayers (1976) conducted an initial series of simple laboratory experiments to gain a feel for the basic response characteristics and reliability of subjects performing a simple task and to investigate various performance shaping factors. For a simple surveillance task requiring a response when an instrument showed an abnormal reading, the probability of a failure to respond

Table 4

ERROR CLASSIFICATION SCHEME

GENERAL CATEGORY	SPECIFIC CATEGORY
1. Observation of System State	a. incomplete b. misinterpreted c. repeated
2. Choice of Hypotheses	a. inconsistent with symptoms b. consistent but unlikely c. consistent but costly d. functionally irrelevant
3. Choice of Procedures	a. incomplete b. inappropriate c. lack
4. Execution of Procedures	a. omission of steps b. other c. inadvertent action
5. Consequence of Previous Error	

Table 4. Analysis and classification of human errors in troubleshooting live aircraft powerplants. (From Johnson and Rouse, 1982)

was found to be 0.001. No effect on reliability of physical environment was observed, within normal control room limits, nor any effect of time on shift (up to 8 hours). However boredom was found to lead to marked deterioration in performance.

Rouse and Rouse (1982) looked at how cognitive style, in terms of personality attributes, affected error rates in simulated fault diagnosis performance. . Two of these personality attributes, felt to be related to problem solving, were examined. A "Matching Familiar Figures" test was used to assess the subjects reflectivity - impulsivity and an "Embedded Figures" test assessed degree of field dependence, (the extent to which one's perception is biased by the configuration of the perceptual field). Performance on two fault diagnosis tasks was examined. The first task was context-free and involved detecting faults in graphically displayed networks. The second task was a car simulator involving troubleshooting mechanical failures. Aircraft maintenance personnel were used as subjects.

The main result was that reflective subjects made significantly fewer errors. The authors note that 'since Henneman (1981) showed errors were one of the three key dimensions of fault diagnosis performance, this result has important implications for selection and training for fault diagnosis. Because there is little evidence that adults can be trained to be reflective, it may be desirable to select inherent reflectives for fault diagnosis training.'

3.7.6 Verbal Reports

Bainbridge (1979) discusses verbal reports as evidence of the process operator's knowledge. She gives evidence showing operator actions and verbal reports are not necessarily correlated and points out that the validity of verbal reports depends closely on the method of collecting them. To minimize distortion in verbal reports, she concludes that a study should be carried out in a social and unconstrained atmosphere, and should access the operator's knowledge appropriately, that is using non-verbal methods to access non-verbal knowledge, and using the operator's language to access verbal knowledge.

Bainbridge emphasizes the absolute necessity of validating reports against non-verbal behaviour. Methodology is reviewed for studies which have assessed knowledge of process control. These studies used written techniques of system state diagrams and questionnaires and verbal reports involving spoken questions and answers or simulation where the operator comments on what he is doing as he does it or verbal protocols where the operator's comments on his behaviour during an actual work situation are tape recorded. Table 5 shows a summary of problems with different techniques and the information which might be obtained from them.

Bisseret (1971) describes methods used to determine the mental processes of air traffic controllers with the aim of automatizing the air traffic system. Many of these methods could be used to analyze the behaviour of nuclear power plant controllers. He also describes variations on some of the verbal report methods discussed by Bainbridge (1979). These include interviews during which a controller solves a real traffic problem and interviews for the same situation in which the controller must ask the experimenter for the necessary information. Content and structure of the operators' memory were assessed by asking them in the middle of a simulation to recall all they knew about the traffic situation being displayed. Recently, Eriksen and Simon (1984) have published a formal theory of protocol analysis, based mainly on laboratory tasks.

Information which might be obtained using different techniques: see text for fuller discussion. (?: not investigated so far, not possible to make a "definite guess")

Technique / Information on:	Questionnaire	Interview	Static simulation	On-line interview	Verbal protocol	Observation
General information on which variable affects which	perhaps	yes	yes	yes	perhaps	infer
General information on control strategy	?	yes	yes	yes	perhaps	infer
Numerical information on control strategy	no	no	no	no	perhaps	yes
Technical aspects of process	?	yes	yes	yes	perhaps	perhaps infer
Decision sequences	no	yes	yes	yes	yes	perhaps infer
General types of cognitive processing	no	?	?	?	yes	no
Full range of behaviours	perhaps	yes	perhaps	no	no	no

Table 5. Information from verbal reports.
(From Bainbridge, 1979)

4.0 PROPERTIES OF HUMAN INFORMATION PROCESSING WHICH TEND TO PRODUCE ERRORS IN COGNITIVE ACTIVITIES

4.1 Human Information Processing and System Design

The following is a brief summary of the better established facts about limitations in human information processing derived from general behavioural science literature. It is the role of the system designer to support efficient operator behaviour by either eliminating from the system any need for behaviour from the system which is sensitive to such limitations, or supporting the operator with appropriate decision aids.

4.2 Memory

1. Not more than about six to eight unrelated items can be held in short term working memory.

2. This span is reduced if some items must be rapidly replaced while others are retained ("running memory span").

3. Memory tends to be poor if a programmed sequence of information processing is interrupted to perform another task involving similar material.

4. Retrieval of information from long term memory by recall is far inferior to retrieval by recognition. (In recognition, an example is available and the person judges whether or not the example matches something in memory.)

4.3 Attention

1. Eye movements constrain visual sampling to less than three samples per second.

2. Fine visual information can only be sampled at the point of regard.

3. Rapid fluctuation of visual attention can be represented as increasing the noise/signal ratio of the observation and hence degrading the quality of data.

4. The occurrence of information about a source in different spatial locations rather than always in the same place increases the difficulty of data acquisition.

4.4 Speed-Accuracy Trade-off

1. The faster a response the less accurate it will be usually be (for muscular actions).

2. The shorter the time during which visual attention is paid to a source the worse is the noise/signal ratio.

3. The shorter the time spent on thinking about data the poorer the quality of the conclusions reached. (This is not well supported by data -- little research is available.)

4. Information presented only briefly is imperfectly encoded.

4.5 Cognitive Tunnel Vision

1. Faced with a problem humans form a plausible hypothesis as rapidly as possible on the basis of as little evidence as possible.

2. Such hypotheses tend to identify the problem as one with which the operator is familiar.

3. Confirming evidence is much more heavily weighted than disconfirming evidence.

4. Only those sources of evidence deemed relevant to confirming the hypothesis are sampled. The rest of the system is often ignored.

5. "Optimistic" hypotheses tend to be accepted more readily than "pessimistic"

4.6 Statistical Inference

1. People undervalue extreme values of statistical distributions.

2. People are poor at combining evidence so as to revise probabilities.

3. People are poor at estimating the range of variation.

4. People require more evidence than is necessary to reach a decision.

5. People think that small samples are more typical of a distribution than is the case.

6. Correlation is frequently interpreted as causality.

7. Subjective extrapolation tends to be linear.

8. Only values and their first derivatives (rate of change) can be perceived without explicit displays.

9. "Regression to the mean" is ignored. People find it difficult to realize that extreme values must be followed by values nearer to the mean, and tend to expect that the occurrence of, say, very high values implies that more high values are to be expedcted. This is incorrect if the parameters of the statistical distribution are constant.

4.7 Logical Inference

1. People find certain rules of logical reasoning difficult. (Wason and Johnson-Laird, 1968).
2. Probable and valuable hypotheses are accepted on slight evidence: costly and improbable hypotheses are rejected or require much evidence.

4.8 Decision Making

1. Adjustment and anchoring -- when overwhelmed by excessive data the observer selects an arbitrary reference point and rescales the data resulting in biased estimates.

2. Availability -- easily available data are given more weight than they deserve and believed to be more probable than they are.

3. Base Rate -- the likelihood of occurence of two (or more) events is estimated by their relative likelihood and the absolute base rate of any is ignored.

4. Data Presentation -- summarized data make a greater impact than raw data; the scaling of graphs has a large effect.

5. Easily recalled data are perceived as more probable.

6. Strongly held values are seen as facts.

7. Statistical fluctuation is perceived as causal.

8. Events are seen as more probable with hindsight.

9. A good outcome due by chance to a poor decision is seen as having been under the decision maker's control.

10. The more data available the more confident observers are in their judgement, irrespective of the quality of the data.

Almost all the above claims are well documented in psychological reports of well controlled laboratory experiments. It is not true that any particular effect will be seen on any particular occasion. And it is important to note that intense practice may lead to cognitive behaviour becoming automatic (performed without conscious effort). When that happens the limitations on operators' abilities are reduced: they become more efficient. Evidence to support the above list will be found in reviews such as Moray (1985), Sage (1981), Johnson-Laird and Wason (1970) and Wason (1968).

In Section 6 this list will be used to develop an empirical model for predicting error rates at the ordinal level of measurement.

One word of warning is appropriate. The generalizations given in this section are frequently based on laboratory, rather than field studies. A safe generalization is that highly experienced operators carrying out real work are likely to be much more efficient than laboratory studies suggest. This section may be overly pessimistic (but correspondingly safer than would be one based on optimism).

5.0 SYNTHESIS

5.1 The Definition of Error

The dictionary (Concise Oxford) is relatively unhelpful. Error is defined as: "mistake; condition or erring in opinion"; mistake, in turn, is defined as: "misunderstanding of a thing's meaning; error, fault, in thought or action." Both appear to have some cognitive involvement implicit in them.

It is more useful to define error in terms of underlying intention of the actor.* Thus if an intention is unfulfilled, an error may be said to have occurred. If the intention had been correct, then the error was one of execution. Similarly, if the wrong intention was formed (analogous to the "erring in opinion") the same action is now the result of a different kind of error. Such an error may be said to have been a cognitive error.

It is clear that the mere detection of an error as manifested in behaviour does not inform us as to the kind of error it was or as to the mechanism which brought it forth. It will require a more intensive analysis to do that.

5.2 Error Taxonomies

Many taxonomies of error have been constructed. Some of these have classified errors as to their effect on the system being controlled. This attitude is the basis for the common tendency to report on and analyze only accidents and to deal only with those errors which were or were assumed to be the cause of the accidents.

Some taxonomies classify errors according to what was done or incorrectly done. Thus, the incorrect turning on of a switch would be grouped with other incorrect turnings on of switches, and so on.

At a more general level, errors can be classified phenomenologically. Kollarits (1937) used a four-fold system: substitution, omission, repetition, and insertion. This system has the advantage of removing the action from the task-specific context and treating the error as a generic behavioural phenomenon. It is our feeling that this is necessary if any kind of general model or theory is to be found.

At a more detailed level, errors may be classified as to the presumed locus of their cause. Thus errors may be either endogenous or exogenous (Senders, 1983); or may be assigned to perceptual, intentional, or action systems in the actor.

---

*The term 'actor' will be used here to designate the person committing an error unless the context makes the term 'operator' more appropriate

5.3 The Question of Whether Human Error is Caused

At the micro level, errors may be considered to be uncaused, or at the most indeterminate, events which arise from molecular or quantal variations in the central nervous system of the actor. Thus, such errors might occur randomly in the same way that alpha particles are emitted by some isotopes of radium.

In prospect, the prediction of error is impossible. One can predict that errors rates will increase or decrease if adequate data have been accumulated about similar situations to that which obtains at the moment of prediction. It is easy to predict that the probability of writing 1984 shortly after New Year's Day, 1985 is high. One might have made the prediction, although slightly less confidently, in the absence of data. Exactly when each of the 1984-1985 errors will occur is not possible to predict. The same limit applies to prediction of NPP operator errors.

In retrospect, on the other hand, it always appears simple to find the cause of of an error by examining the sequence of states of plant and operator which preceded the error. Thus, errors which have not yet happened might be considered as uncaused and those which have already happened as caused. It is probably scientifically sounder to accept the idea that errors are uncaused. To assume cause can lead one into a philosophical morass and does not necessarily lead to ways of eliminating the errors. At the same time one may well be able to state what factors or situations cause errors to be more probable (cause error rates to rise).

In brief, looking forward in time, one sees behaviour as being stochastic in nature, and one sees error as the accidental outcome of the sequence of events not truly under the control of the actor. Looked at backwards in time, one can trace any particular outcome to the antecedent elementary states of the system and assign cause to what has been called an error. Thus errors are simultaneously caused and uncaused.

5.4 On the Modelling of Human Error

5.4.1 Levels of Analysis

It is convenient to distinguish many aspects of human error. The following table presents a way of distinguishing the causes and modes of error from the expressions and consequences.

Possible Causes	Taxonomies				
	Locus of Error	Process Involved	Mode of Error (Phenomenology)	Expression	Consequence
Common Element	Endogenous	Perception	Substitution	These depend on what was to be done and on what could be done and are highly plant and layout specific.	These depend completely on the nature of the plant and its state at the time of the error.
Loss of Intention	Exogenous	Decision	Ommision Insertion		
Lack of Time		Intention	Repetition		
Ignorance		Choice	Slips		
Forgetting		Execution	Mistakes		
Random Neural Events					

The descriptions in Column 1 attribute error to mechanisms internal to the actor. These vary widely. Thus errors may be considered to be the result of random uncaused events in the central nervous system of the actor. Such errors might be presumed to occur as spontaneously as the emission of an alpha particle from a radium isotope. Alternatively they might be attributed to hypothetical micro events at the neural level as opposed to the quantal level. Random variations in sensitivity of neurons to near-threshold signals may give rise to random errors in behaviour.

At a higher level of attribution, errors (capture, for example) might arise as a result of similarities in the components of a cognitive task.

In column 2, errors are classified as to whether they are the result of internal or external mechanisms, in column 3 according to the presumed locus within the actor of some cause, and in column 4 by their manifestation in behaviour.

Column 5 tabulates errors according to what actually was done by the actor. Thus opening a valve (or leaving it open instead of closing it) would be one category. The layout of the control room obviously limits what can be done wrongly.

Finally, column 6 classifies errors according to what happened to the plant. An error might appear in this column in one plant state and not in another although both occurred. Neither of the schemes in column 3 or column 4 would be concerned with error if there were not something that happened in the plant context.

Our proposed model and exhaustive classification scheme offers a taxonomy which subsumes most of the other schemes. It points out very clearly that executions or actions do not allow one to designate mechanisms or causes, and further, that expressions and consequences are wholly dependent on systems and workspaces.

It is our view that the principal concern should be with columns 2, 3, and 4 if prediction is the goal. It is possible that research should be carried out on the topics of column 1, if prevention is the goal, in order to improve our understanding of the mechanisms which give rise to the errors in column 2. To pay too much attention to the contents and classifications of columns 5 and 6 would lead one into the position of fighting brushfires on a continuous basis with no possibility of achieving general solutions. Unfortunately most of the operational data on human failure fall into categories 5 and 6.

The modes of error may be limited in number whereas the expressions of error are essentially unlimited. A reporting scheme should at the very least adopt a standard classification scheme and be used for all instances of human failure.

5.4.2 An Exhaustive Model of Human Error: EXHUME

It is also convenient and useful to consider the disposition of errors. If we consider that within the actor there are a series of information processing modules, then we can consider the probabilistic distribution of "signals" passing through the series.

The following figure (Figure 12) assumes that some plant state is available to the actor (at the left end of the chain) and is processed by various components in turn. At each process, there is a probability that an error will be made. When an error is made, there is a probability that it will be self-detected by the actor and the information processing stage will repeat. There will be once again a probability of an error, and so on. Also, the case that an error can be followed by a compensating error leading the sequence back into the path of correct processing. The actor is bounded at both ends by the transition, at the left, from the plant state to sensory machinery, and at the right, from mode of error to expression of error.

Although many of the probabilities designated are possibly not within reach of experimental method, some of them, at any rate, can be estimated by inference from those which can be measured. At the very least, the model is a useful conceptual tool for thinking about human error.

The cognitive functions of the operator of the complex system are those which involve perception, choice of goal and choice of action. To the right, the last 2 boxes contain the consequences, and the expressions, of errors. These plant and task-specific functions form the usual content of the L.E.R.'s and S.E.R.'s.

Each link is identified with a probability which designates the source and the sink of the link. Thus P(1,20), a failure to observe, occurs because of a signal to noise problem. P(1,21) is a miss which occurs because of a lack of observation or attention.

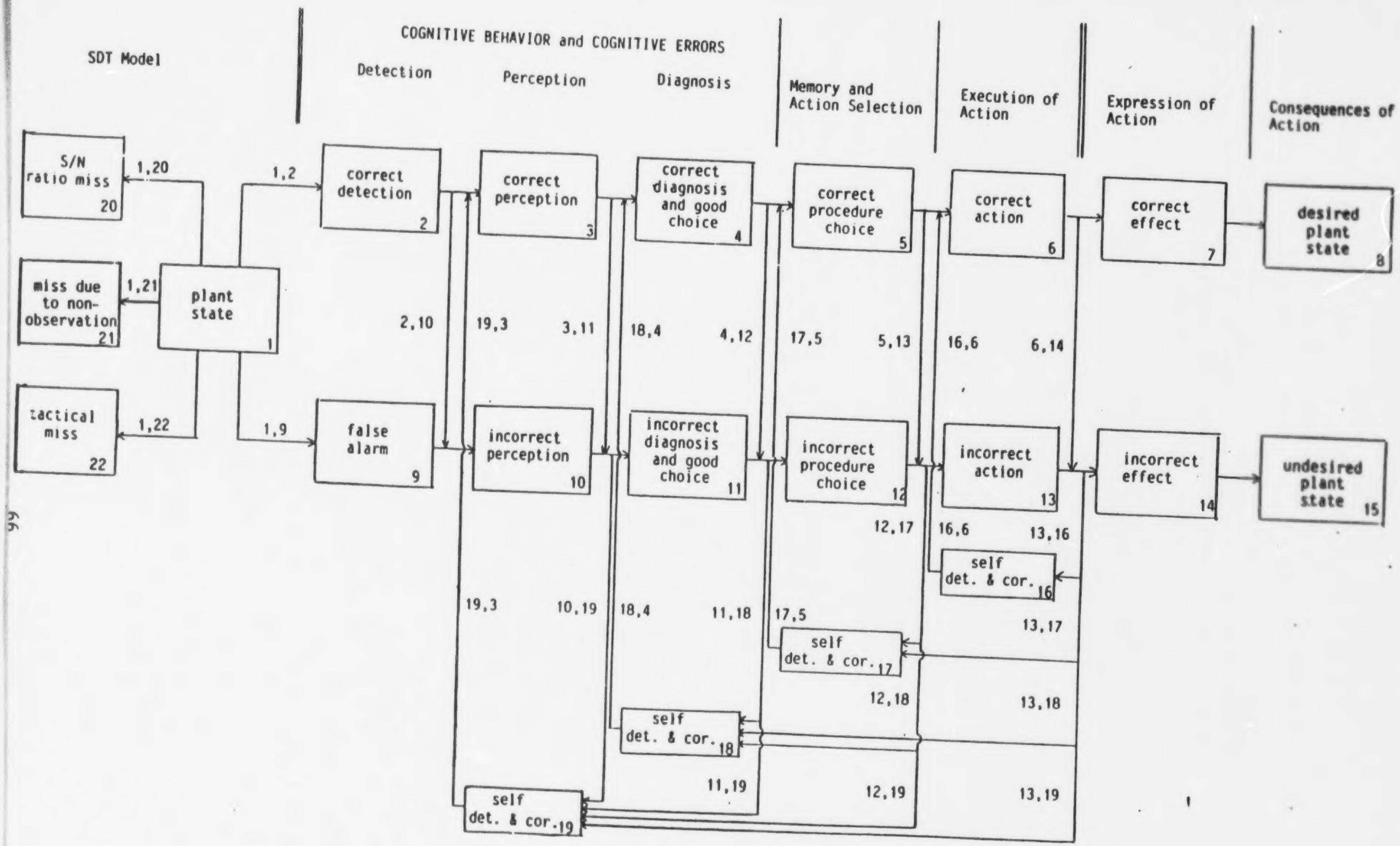

SDT Model
COGNITIVE BEHAVIOR and COGNITIVE ERRORS
Detection
Perception
Diagnosis
Memory and Action Selection
Execution of Action
Expression of Action
Consequences of Action
S/N ratio miss 20
miss due to non-observation 21
tactical miss 22
plant state 1
1,20
1,21
1,22
1,2
1,9
correct detection 2
false alarm 9
2,10
correct perception 3
incorrect perception 10
19,3
3,11
10,19
correct diagnosis and good choice 4
incorrect diagnosis and good choice 11
18,4
4,12
11,18
11,19
correct procedure choice 5
incorrect procedure choice 12
17,5
5,13
12,17
12,18
12,19
correct action 6
incorrect action 13
16,6
6,14
13,16
13,17
13,18
13,19
self det. & cor. 16
self det. & cor. 17
self det. & cor. 18
self det. & cor. 19
correct effect 7
incorrect effect 14
desired plant state 8
undesired plant state 15

Figure 12. EXHUME: A Multi-Staged Closed Loop Model of Behaviour

P(1,22) is an error which occurs because the observer's criterion is too high resulting in a "tactical miss".

It is profitable to assign names to the various links, both forward and backwards as follows:

Link Number	Error type
1,20	miss error due to low S/N ratio
1,9	false alarm due to choice of decision criterion
1,22	miss error due to choice of decision criterion
2,10	perceptual error
9,10	perceptual error stemming from false alarm
3,11	diagnostic error
10,11	diagnostic error stemming from perceptual error
4,12	procedural error
11,12	procedural error stemming from diagnostic error
5,13	execution error
12,13	execution error stemming from procedural error

Naturally the remaining forward links are all correct performances. Thus:

1,2	correct detection or "hit"
2,3	correct perception of plant displays
3,4	correct diagnosis of plant state
4,5	correct choice of procedure
5,6	correct execution of planned actions

The backward running links are all the mechanisms by which self-detection and correction (sdc) of each of the error types can be accomplished. Thus:

Link Number	Error Type
10,3	sdc of perceptual error
11,3	sdc of diagnostic error stemming from perceptual error
12,3	sdc of procedural error stemming from perceptual error
13,3	sdc of execution error stemming from perceptual error
11,4	sdc of diagnostic error
12,4	sdc of procedural error stemming from diagnostic error
13,4	sdc of execution error stemming from diagnostic error
12,5	sdc of procedural error
13,5	sdc of execution error stemming from procedural error
13,6	sdc of execution error

Each of the various components of the model may conveniently be named as shown in Figure 12. We have therefore:

Detection
Perception
Diagnosis
Action Selection
Action Execution

and for most of these suitable micro models are available.

Detection can be successfully modelled as a Signal Detection Theoretic process, SDT; Perception as a pattern recognition process coupled with a SDT model; Diagnosis again as either SDT or Information Theoretic model; Action Selection as a serial process model and Action Execution as a serial process model.

In all cases and at all levels of description and modelling, it is important to remember that the errors arise within the actor. How they are in fact expressed to the outside world is a function of what the actor is trying to accomplish.

5.5 The Problem of Detecting Cognitive Error

Many of the elements that we have suggested make up cognitive behaviour in the context of a NPP are most easily modelled as stochastic processes. As hinted earlier, SDT is the model of choice. A consequence of this makes difficulties for the analyst. A correct diagnostic choice does not always in fact lead to a correct outcome. Whether it does depends on the actor's choice of likelihood ratio for false alarms and misses. This in turn depends on the pay-off structure under which the actor is working. The decision not to wear a seat-belt may in fact lead to an escape from an automobile fire even though it is the incorrect strategy for avoiding injury in automobile accidents. The presence of an undesired outcome does not necessarily imply error; neither does the presence of a desired outcome imply the absence of error. Errors followed by a desired outcome can be detected only through protocol analysis or through the analysis of large numbers of data from which the decision strategy of the actor can be established and compared with some ideal decision strategy. This means, in turn, that data gathered from operating NPP events do not inform us usefully about the presence or absence of operator diagnostic errors. A different kind of data must be gathered. Even then it is not clear that enough events have occurred or will occur to enable us to carry out an analysis.

One can see two alternatives. In-depth debriefings of operators after each event can yield protocols which may be useful in determining whether an actual cognitive error preceded the unwanted outcome. Or a laboratory can be set up, designed to permit the gathering of adequate data about decision making in the diagnosis of events in complex plants. It is conceivable that a dedicated NPP simulator could be used for this purpose. In all probability it would require more than the usual endowment of flexibility of control and memory. It would probably be cheaper and more effective to design and build a research facility specifically for the study of human error under experimenter control and with sufficient flexibility to permit easy experimenter variation of any of the multitude of variables which might influence cognitive behaviour.

5.6 Multi-Stage Closed Loop Model of Behaviour

Nowhere in the literature have we found a model or a taxonomy which is satisfactory for the identification of cognitive errors and for guidance as to the remedy of them. The systems which are aimed at providing a data base for probabilistic risk analysis, for example SLIM-MAUD, do not provide a basis for distinguishing between cognitive and non-cognitive errors. In general, none of the taxonomies or models which we encountered stands up to rigorous examination. The psychological models, that of Norman for example, do not provide a basis for quantification.

Our new model, EXHUME, presented above in section 5.4.2, clearly identifies the kinds of error which can occur and, as will be shown, provides a basis for an experimental method for obtaining the necessary data for quantification and prediction. It was designed for that purpose. It is both a model of error production and a multilevel taxonomic scheme. It points out that the various modes of error and the various expressions may arise from widely differing mechanisms. The collection of operational data about modes, and even less data about expressions, will do little to inform us about underlying possible cognitive errors, and the collection of data about expression of error even less.

The first level of the model considers the detection process itself. The plant state, at the left, is presumed to exist as a signal. The nature of the signal is plant and layout specific. Signal Detection Theory provides a basis for quantification of the behaviour of the actor at this point in the flow diagram. The signal may have a signal/noise ratio which permits easy discrimination between signal and noise. If the actor fails, however, to detect the signal (a miss) there will be an error, which may be an omission, an insertion or a substitution (although hardly a repetition) since a moment's thought shows that a missed signal may give rise to a false diagnosis which in turn gives rise, at the right, to any of a variety of modes of error. The expression of the resulting mode will depend on what is available to be done in the specific context of the workspace.

The signal/noise ratio which will give rise to an error is design specific. Any signal could have been made strong enough so that the probability of its being missed is vanishingly small on the grounds of signal strength alone. A s/n error at this point is a non-cognitive error.

An error at the input stage may also arise as a result of the operator's choice of decision criterion. The choice of decision criterion can vary for any of a variety of reasons. Fatigue, training, motivation and labour-management relations can all influence an operator's choice; the last of these may well be the most important. These are cognitive errors.

The next stage in the process shown in Figure 12 is clearly in the cognitive domain. Here a perceptual process is involved in transforming the detected signal into a recognized plant state. No 'micro-model' comes readily to mind which can help in the analysis of this stage. It may be that pattern recognition schemes (in the artificial intelligence literature) will provide a basis for predicting the probability of correct and timely identification of plant states from the detected signals.

The diagnosis stage appears to be analyzable by means of Signal Detection Theory. Diagnosis, as has been earlier remarked, is a statistical process and correct decisions are not necessarily followed by correct results. A strategically correct decision criterion which derives from an intense concern for safety may yield a great many false alarms. One based on a production-maximizing philosophy may result in exceeding nominal safety margins on many occasions as a result of missed signals.

The stage of procedure choice should not be cognitive. If a decision has been made about plant state and a goal chosen, the procedure should be either a reliable automatic choice based on extensive training and operational experience, or a fixed look-up procedure to be followed no matter the level of confidence of the operator. If, on the other hand, the operator is left to his own devices as to choice of procedure, such choice would be cognitive and erroneous choices would be cognitive errors.

If we eliminate this latter from consideration, we are left with three sources of cognitive error for which models may exist for quantification and the guidance of experimental investigation. These are: Detection, Perception, and Diagnosis.

5.7 The Information Required for Quantitative Predictions

The eleven kinds of error were identified in Figure 12. Each of these has a probability attached to it. The problem is to determine or estimate the value of this probability. In particular we wish to have the values for the following flow lines:

1,9	false alarms due to inappropriately low decision criterion
1,22	misses due to inappropriately high decision criterion
2,10	failure to recognize plant state
3,11	incorrect diagnosis of plant state

and the values for self-detection and correction for the following:

10,19,3	sdc of perceptual error 2,0
11,19,3 }	
12,19,3 }	all of which stem from perceptual error 2,10
13,19,3 }	
11,18,4	sdc of diagnostic error 3,11
12,18,4 }	
13,18,4 }	both of which stem from diagnostic error 3,11

Whether we use expert opinion or experimental evaluation or the in-depth analysis of debriefing protocols from operational experience, what we should seek are the numerical values of these probabilities.

5.8 Error Model Based on Empirical Data (EMBED)

EXHUME is an analytic model. It is based on a logical analysis of the nature of information processing required by different tasks. It is however possible to approach error prediction in another way, using empirical data on how operators process information rather than a logical analysis.

In Section 4 of this report a list was given of 38 properties of tasks or people which make it difficult for humans to process information efficiently. That list was derived from a study of research reports published in human factors, ergonomics and general psychological literature. The status of the facts is entirely empirical: there is no theoretical reason why running memory need be less than static memory; there is no theoretical reason why variances should be underestimated; there is no theoretical reason why a decision maker favour seeks confirming evidence rather than conflicting evidence. If is true however that humans generally behave in such ways.

It is possible to construct a model for the prediction of errors by combining these facts. The status of such a model is entirely dependent on the quality of the research and data on which it is based. Unlike an analytic or logical model there can be no guarantee of internal consistency, and no guarantee of completeness. With that proviso, a model to predict the probability of error at least on an ordinal scale ("error will increase", "error will decrease") can be proposed. The model consists of the following steps:

1. Perform a task analysis to define the mental operations involved in each step of a task.

2. Consider these task elements in the context in which the task will be performed.

3. Beside each task element write the number of all entries in the Section 4 list which are relevant to performing that task.

4. Axiom: The probability of error increases as the number of Section 4 elements increases.

As an example, compare two hypothetical tasks. The first is to interrupt an ongoing task to read three gauges and decide whether at least two of them have the same, normal, value. Since at some precision this will not be the case, the task involves looking at each in turn, observing the value, remembering the value while looking at the others, comparing the values of each with the others, and making a decision. In terms of the Section 4 list, the following six items are relevant:

4.2.2 Some items in memory rapidly changed while others are retained for later use.

4.3.2 Fine visual information can be sampled only at the point of regard.

4.3.3 Rapid fluctuation of visual attention (increased the noise/signal ratio).

4.5.5 "Optimistic" hypotheses are accepted more readily that pessimistic ones.

4.8.3 The likelihood of occurrence of two or more events is estimated by their relative likelihood and their absolute base rate ignored.

4.8.11 An observation which is expected is more likely to be observed and an unexpected one missed.

On the other hand, if diagnosis of plant state is involved, the operator will probably have to compare two or more readings, form a hypothesis, decide what procedure to follow, take more readings, remember some while observing others, perhaps estimate the mean and range of a fluctuating value, etc. The list will be much longer, and the probability of a cognitive error much greater.

Graphically, the model can be represented in Figure 13.

In an attempt to assess how realistic would be the proposal to analyze NPP tasks in this way, several EOPs for CANDU reactors were examined. Such written EOPs leave very little room for cognitive action by the operator; the majority direct his attention to a particular display and state the value which is to be expected under various hypotheses. Hence most of the steps to be taken are simple "observe-read-note-decide" sequences. But it may be that there is a greater load on the operator than appears: if two displays were far apart, memory, visual attention, and speed/accuracy factors could be more relevant in some tasks that others. But this cannot be decided by reading EOPs out of context. It is certainly the case that many U.S. EOPs still contain far more cognitively oriented commands, such as "if the pressure is falling too quickly take steps to ensure that the range is reduced." This throws a considerable cognitive load on the operator.

More typical of CANDU EOPs are instructions like the following (which are based upon existing EOPs but are not actually real EOPs).

1. "Determine whether the process computer is running."

a) If there is a specific light, readily distinguishable, which is illuminated if and only if the computer is running, then the major error producing characteristics will be 4.3.1, 4.3.2, 4.3.3, 4.4.2. The task is primarily perceptual.

b) If the operator has strong reason to wish to believe the computer is running, add 4.5.3, 4.7.2.

c) If there is no specific RUN light, it may be that the only way to tell is to try to detect flickering in the accumulator status bit lights: add 4.6.8.

2. "Signs that moderator temperature may be improperly controlled include (1) excessive reduction in level of moderator..."

A rate has to be estimated (4.6.8, 4.6.2, 4.6.5) in a particular subsystem (4.5.4) and the operator has to decide whether the rate is excessive. This involves comparing values over time (4.2.2) while monitoring the entire system (4.3.1, 4.3.2, 4.3.3, 4.4.2). If no explicit criterion is displayed, "excessive" must be judged against a remembered value (4.2.4).

3. "Can the pump be repaired in 20 minutes? If so, initiate repairs. If not, initiate manual shut down."

This task involves assessing the state of the pump, and forming a hypothesis about the cause of its failure (4.5.1, 4.5.2, 4.5.3, 4.5.4). It also involves assessing the overall state of the plant, since to repair the pump involves taking it out of service, and to complete such an assessment many displays must be examined (4.3.1, 4.3.2, 4.3.3, 4.4.2, 4.4.3, 4.4.4).

Figure 13

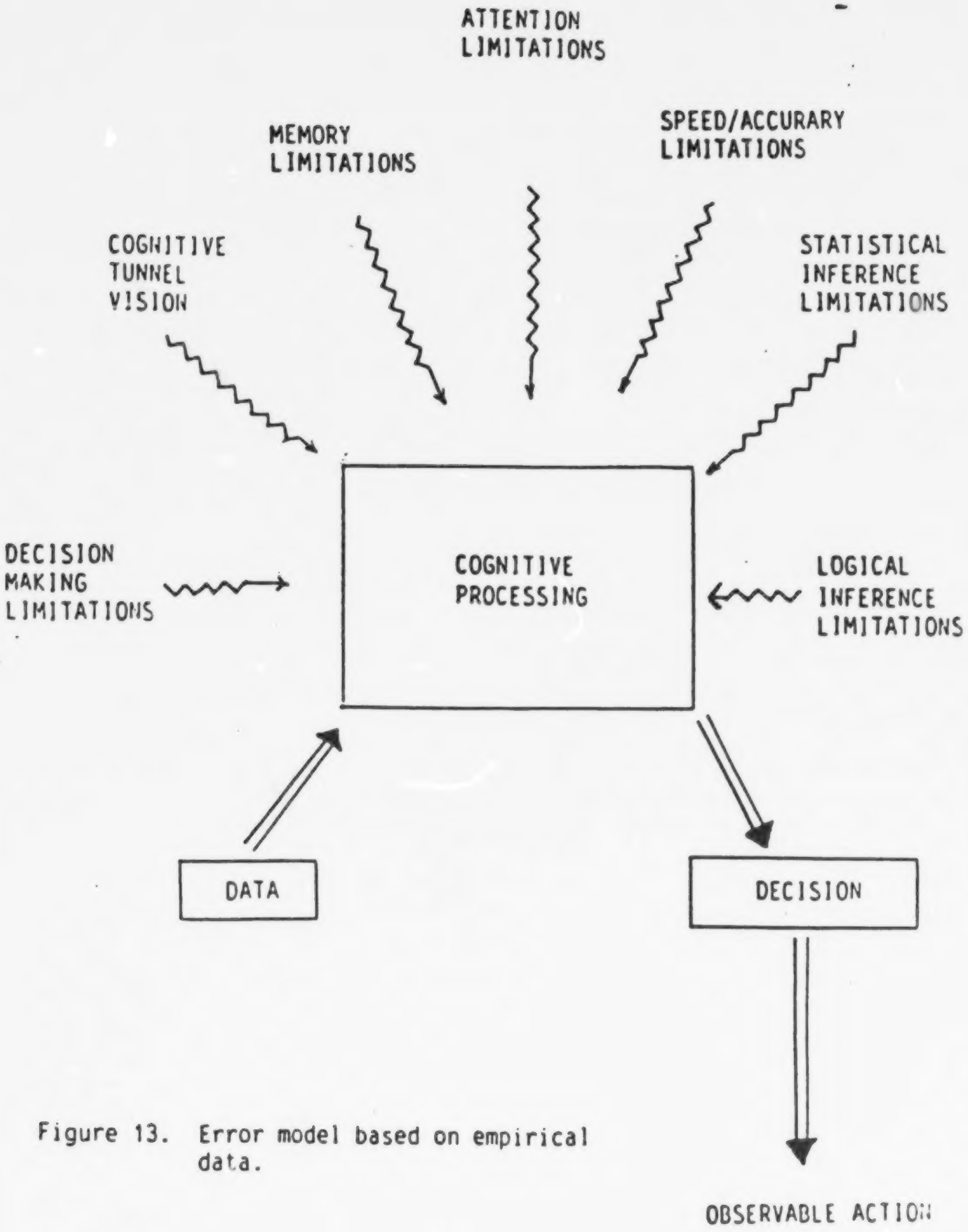

ATTENTION LIMITATIONS
MEMORY LIMITATIONS
SPEED/ACCURARY LIMITATIONS
COGNITIVE TUNNEL VISION
STATISTICAL INFERENCE LIMITATIONS
DECISION MAKING LIMITATIONS
COGNITIVE PROCESSING
LOGICAL INFERENCE LIMITATIONS
DATA
DECISION
OBSERVABLE ACTION

Figure 13. Error model based on empirical data.

The implication of taking the pump out of service must be assessed (4.7.1, 4.7.2). And the symptoms of the pump must be examined in the light of experience to decide what is the nature of the failure (4.2.4) and to decide whether 20 minutes is the usual time which maintenance takes (4.2.4). Finally, communication must establish whether maintenance men are available (variable not in model).

Comparing the three tasks, the number of elements are

Task 3 - 14
Task 2 - 10
Task 1 - 4 to 7

We would therefore predict that the probability of an operator making an error is also in the rank order Task 3 > Task 2 > Task 1.

It may be possible in certain cases to proceed further. For example there are several quantitative "micro-models" which have been validated experimentally. Signal Detection Theory provides a quantitative model for perceptual judgement and the biasing effects of subjective expectancies and payoffs. Optimal Estimation Theory provides an alternative model, which can also be related to speed/accuracy tradeoff and attention; while Optimal Control Theory can be used to compute the final decisions about action in certain kinds of manual control tasks in the time domain. Several quantitative models of visual scanning and sampling behaviour exist, and have been at least partially validated. At a purely empirical level it would be possible to estimate from the literature by what proportion variance is underestimated and by what proportion Bayesian cumulative probability adjustments are underestimated. Some speed/accuracy curves are available, although the degree to which they can be generalized beyond the setting in which they were obtained is doubtful.

Overall, a number of quantitative estimates could be added to the Section 4 list, which would increase the strength of measurement of EMBED in the direction at least of interval if not ratio measurement. Some gaps would still remain to be filled with expert judgement pending further research. The direction and work required to improve EMBED is clear, and would provide converging evidence which would support estimates formed from the quasi-analytic model supported by expert judgement.

6.0 RESEARCH PLAN

6.1 Alternative Ways of Gathering Data

Most of the quantities which are needed even for an evaluation of the importance of the problem of cognitive error are not readily available in the literature or in data stores which may exist. The necessary data must be obtained by establishing methods based on data collection from operational plants or simulators, from expert opinion appropriately extracted, or from experiments performed on apparatus designed for the purpose.

6.1.1 L.E.R.'s and S.E.R.'s.

These reports are predominantly concerned with the expressions and consequences of error. What happened in the plant (or the simulator) is uppermost in the minds of those who are responsible for gathering data. Thus there is much information about what happened and none about why. If an operator is asked about why an error was made (and only such errors which have lead to consequences normally are inquired about) what he gives are reasons rather than causes. Operators are not trained introspectionists and introspectionists are not trained operators. Third party inquiry is required. Events in the plant, if they are to provide a data base for the detection and analysis of cognitive errors must be followed by in-depth interviews by specialists trained in the techniques of interviewing. (Interviews by operators will probably impair the quality of the data from the point of view of error analysis.) Even then it is not clear that reason and cause can be adequately separated. Further, the number of events which might yield useful data will, fortunately, be small. Plant and simulator data may well serve primarily as a validation base for estimates obtained in other ways.

6.1.2 Expert Opinion

The SLIM approach (Embrey, 1983) is a good example of this technique: "Basically the SLIM approach assumes that expert judges perceive the likelihood of successful completion of a task as a function of a number of attributes or 'Performance Shaping Factors' (PSF) such as the quality of information available, competence, team structure, etc." Embrey goes on to claim that the Success Likelihood Index (SLI) works when applied to field situations. His confidence seems somewhat unjustified since no actuarial data were available to calibrate the SLI estimates. Even if such data were available, "... the SLIM technique does not explicitly break down the task into elementary failure modes." Thus the identification of a failure as stemming from a cognitive error would be difficult if not impossible.

In Section 3 of this report several currently available methods for estimating error rates and error possibilities were examined. All of them relied heavily on subjective estimates by subject matter experts. Such methods are well established and widely used in the assessment of the contribution of human error to PRA. But it is well known that there is inadequate empirical evidence to validate the estimates.

It would certainly be possible to obtain subjective estimates of the probability of error at each step in the flow chart of Figure 12. But the validity of the estimates would be even lower for cognitive errors than for behavioural errors. At least in the case of behavioural errors a person sees not merely his own but other people's errors. In the case of cognitive errors far fewer of the errors of others are identifiable because it is unclear whether erroneous behaviour is due to a correct cognitive decision followed by incorrect action, or to a cognitive error followed by action which is appropriate given the erroneous decision. Furthermore, it is not known how often people detect and correct their own cognitive errors.

It seems safer to make use of sdc estimates indirectly, and as far as possible merely to fill in the gap in a logical or empirical model. For example we have discussed elsewhere how error rates at both the front end and output end of the flow chart can be estimated. It would be appropriate to use subjective estimates to put quantitative bounds on the intervening blocks until such time as data are available.

Similarly, in the case of the Empirical Model EMBED, certain quantitative values can be obtained from the literature. For example, research on humans as intuitive statisticians could provide an estimate of the magnitude of variance underestimation, or how far humans fall short of Bayesian probability revisions. On the other hand, subjective estimates are necessary at present for the probability of and duration of "cognitive lock-up". If Embrey's SLIM-MAUD methodology could be validated it would be a particularly suitable approach; but any standard subjective scaling approach could be used.

However, such estimates are not a firm basis for prediction. It must always be born in mind that the criteria for decisions fluctuate from moment to moment due to the interaction between observation, prediction, and action. Humans are dynamic decision makers, and the probability of error consequently fluctuates from moment to moment.

6.2 Research Plan for Error Data Collection

A carefully designed laboratory task of sufficient complexity can be designed. Such a task must be comparable to the tasks which confront the operator of a NPP control room. The difference between such a facility and a NPP simulator is that the facility would be designed to permit the manipulation of task variables in accord with theory to elicit the desired data. Such a facility would require more than the usual endowment of computer memory and flexibility for control of a simulator.

In general, we feel that many, if not all, of the desired probabilities can be obtained through carefully designed and carried out behavioural experiments.

An experimental facility for the investigation of error must be very much like a real work place. We have no good basis for assuming that the errors made in a palpably synthetic work situation will be in frequency or genesis like those which occur in a real workplace. This means that data must be gathered from persons whose job it is to operate the experimental "plant", and on an 8 hour per day, 5 days per week basis. They must be performing "real work".

The experimental facility must be a well human engineered one, with good signals, understandable coherent structure, and suitable training provided. Errors should be no more frequent, on the average, than those in a NPP. This specification leads to a requirement for the experiments to be run for a long period of time in order that a data base can be generated. If we assume that well-trained operators make one error every 10 minutes (as has been estimated for aircraft pilots under stressful conditions) and if there are as few as 11 different kinds of error (according to EXHUME) and if all were of the same frequency of occurrence, only one data point per error type per 100 minutes would be obtained. If a sample of 100 errors of each type were desired, we would have to run the experiment for 11,000 minutes or twenty-four 7.5 hour days, per person. If errors actually occur only every 100 minutes, the experimental plan would call for 244 days of operation - a full year (assuming the usual holidays and sick leave and so on). This latter seems like a more reasonable estimate. Although it is not a welcome conclusion, one must remember that it takes time to investigate relatively rare phenomena.

We strongly recommend that such long term studies be considered. In the long run, we feel that it is the only way for definitive data to be obtained and values assigned to the various probabilities of cognitive error.

Once such data have been gathered, it will be possible, (again over non-trivial periods of time,) to explore the effects of remedial processes applied to the operators and the workstations.

6.3 Research Plan for EMBED

The following plan would lead to a clear picture of the extent to which empirical data can at present support estimates of error based on a causal model of the psychological mechanism involved.

For each item on the Section 4 list the behavioural science literature should be searched for relevant quantitative data.

In the case of items for which such data exist an estimate should be made of the confidence intervals available. Careful support or rejection of the validity of transferring the data from research to field settings should be made. Special attention should be paid to problems of advanced control rooms.

Where possible, the empirical data should be embedded in a general quantitative model of the cause-effect relation.

Where gaps are found which prevent quantitative estimates of error probabilities, subjective expert judgement should be temporarily used, with the acknowledgement that in all such cases the result is probably to drop the precision of measurement from interval to ordinal.

A program of experimental research should be mounted to plug such gaps. While this can be carried out in the laboratory it is important that complex rather that simple tasks be used in order to highlight interactions, and operators well protected should be trained.

A model should then be constructed on the assumption that the proposed causal mechanisms will generalize to cognitive processing because the latter, occurring in the head, is to a considerable extent independent of specific plant characteristics.

The model should be compared with expert judgement, the flowchart model, and such empirical field data as exist.

The two research proposals discussed in sections 6.2 and 6.3 are complementary.

6.4 Research Plan for Quantification of Cognitive Error Mechanisms.

The nature of the model EXHUME is such that experiments can easily be devised to test some of the links between the boxes; thus the link connecting plant state and tactical miss can be identified and quantified using the classic signal detection paradigm. The link connecting correct choice of action and incorrect excecution can be quantified by operator interviews and post-error protocol analysis. Cognitive errors, which are presumed to occur within the operator, will be slightly more difficult to identify and quantify directly, but we believe that a program can be devised to obtain most, if not all of the desired quantities.

Cognitive error manifests itself as an incorrect action. The task is to identify the failure link, one of the vertically oriented lines in Figure 12 which induced the error. Let us imagine that an incorrect diagnosis was made. There is a finite probability that a self detection will be made and a new diagnosis performed. This is an event which the operators themselves can be instructed to report. If the erroneous diagnosis is not identified and corrected, it will result in an incorrect choice of procedure, an incorrect selection of action, and an excecution of the incorrect action. (The second and higher order probabilities are vanishingly small, and can be ignored.) If the operator makes an incorrect action, the manifest or expressed error, a simple question will reveal what procedure was chosen which led to the action. If the procedure chosen was the correct one, then the error will have been an incorrect choice of action. This too, will be revealed by questioning of the operator. Finally, if the correct action was chosen, the error must have been non-cognitive execution error.

In general, the mechanism of the error should be evident to the experimenter, even if not to the operator. The operator might even be asked to identify the procedure chosen when it is chosen, by following his own steps in an overt way.

A suitable measure of attention would also greatly simplify the identification of error mechanism. The eye fixation of an operator will help us to identify the nature of the information which he seeks, and therefore the nature of the internal model which guides the operator's behaviour. We believe that methods can be developed to identify the point of regard as a function of time of each of the operators in a complex multi-person supervisory control task. We further feel that such measurements are neccessary for a complete understanding of operator cognitive behaviour.

7.0 CONCLUSIONS

1. At present it is not possible to predict the probability of cognitive errors, or to estimate cognitive error rates at more than an ordinal level. That is, one can specify conditions under which errors are more or less likely. One cannot put a reliable number in the range (0,1) on the probability that a cognitive error will occur

2. Since decision criteria are dynamic, any estimate must be regarded as non-stationary.

3. In principle, models are possible based on observed error data, on logical/analytic concepts, and on general empirical principles of behavioural science. In practice at present only the last two seem practical. Both would have to include some elements of expert judgement in the estimates. The safest claim would be to ordinal measurement in the foreseeable future. Empirically derived models of time-to-begin or time-to-complete actions have been introduced in some U.S. PRAs. The problem is that these items can be very sensitive to context. For example, time estimated from LERs are about 8 times as long as times in a training simulator, and until more is known about what factors change the times it is unsafe to rely on these estimates.

4. There may be special problems associated with advanced control rooms which will increase the probability of cognitive errors relative to action errors, and may indeed increase the absolute level of cognitive errors.

5. EXHUME is a basis for research. It identifies the kind of data required for a comprehensive analysis of error. Careful experimental design will make it possible to gather the data. This model also makes clear the relationship between cognitive error and the kind of data available in L.E.R.'s and S.E.R.'s.

6. The EMBED model requires detailed research to identify exactly which empirical findings are most reliable and most directly relevant to error prediction. It should be possible to use EMBED at least to estimate relative probabilities of errors for different tasks, and at least in some cases to make absolute predictions.

8.0 ANNOTATED BIBLIOGRAPHY

The following is a bibliographic list of the publications collected and selected as relevant to the project. Each reference is briefly described.

Abo El-Ela, M.A., and Raouf, A., 1979. Manual backup operations: Some behavioural aspects of human reliability. Microelectronic Reliability., Vol. 19, 141-149.

A simulated industrial task was used to investigate the factors affecting reliability in manual backup operations. A model is presented.

Adams, J.A., 1982. Issues in human reliability. Human Factors, 24:1,. 1-10.

Human reliability cannot be expressed in equipment reliability terms. Methodological problems are due to lack of a definition for human failure and units of behaviour. There is a lack of methods to synthesize the units to determine behaviour sequence reliability and to integrate human with equipment reliability. Suggests Monte Carlo modelling. (Cites Fleishman's 1975 taxonomy of 4 kinds of conceptual approaches to classifying the units of behaviour.)

Altman, J.W., 1967. Classification of human error. In Symposium on reliability of human performance in work.. W.B. Askren (Ed.). 1966 Annual Convention of the American Psychological Association, May 1967. AMRL-TR-67-88.

Classification of error, error analysis at different levels.

Anyakora, S.N., and Lees, F.P., 1974. Detection of instrument malfunction by the process operator. In E. Edwards and F.P. Lees (Eds.), The Human Operator in Process Control. Taylor and Francis Ltd. London, 238-248.

Chemical plants: Malfunctions -- definition, detection, emergencies.

Askren, W.B., and Regulinski, T.L., 1969. Quantifying human performance for reliability analysis of systems. Human Factors, 11(4), 393-396.

General model for time-continuous tasks tested in the lab. Errors fit Weibull, gamma, and log-normal distributions.

Bainbridge, L., 1969. Verbal reports as evidence of the operator's knowledge. International Journal of Man-Machine Studies. 11, 411-436.

A discussion of methods for collecting verbal reports for inferring operator's knowledge. They are analyzed in terms of the correlation between the reports and observed behaviour.

Beare, A.N, Dorris, R.E., Bovell, C.R., Crowe, D.S., and Kozinsky, E.J., 1984. A simulator-based study of human errors in nuclear power plant control room tasks. NUREG 3309, Jan.

Quantification of error rates during normal operations and a comoparison to the results in NUREG 1278. Found the estimates to be largely in agreement.

Bell, B.J., and Swain, A.D., 1981. A procedure for conducting a human reliability analysis for nuclear power plants. NUREG 2254, Dec.

A step-by-step procedure to estimate probabilities of failure due to human error.

Bisseret, A., 1971. Analysis of mental processes involved in air traffic control. Ergonomics, Vol. 14(5), 565-570.

Characterizes the internal operations of ATC controllers as a categorization task and tests this hypothesis through interviews with controllers given ATC problems to solve.

Braune, R., and Forshay, W.R., 1983. Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. Instructional Science 12, 121-145.

Relationship between problem-solving, expectancies and schemata. Cognitive interactions should be modelled as levels of judgement and decision making. Operators may fail to attach the appropriate meaning to stimuli in the environment.

Braune, R.J., and Trollip, S.R., 1982. Toward an internal model in pilot training. Aviation, Space, and Environmental Medicine, Vol. 53(10), Oct. pp. 996-999.

Training mostly but discusses the importance and role of internal models (e.g. defines comprehension,; prediction). An analysis of phase-plane switching loci. Introduces a method for slow systems based on the hypothesis that sub-optimal performance originates from the operator's inability to predict system motion.

Comer, M.K., Kozinsky, E.J., Eckel, J.S., and Miller, D.P., 1982. Human reliability data bank for nuclear power plant operations. Vol 2: NUREG 2744.

de Jong, J.J. and Koster, E.P., 1974. The human operator in the computer-controlled refinery. In: E. Edwards and F.P. Lees (Eds.). The Human Operator in Process Control. Taylor and Francis Ltd., London., 196-205.

Discussion of the operator's role under normal and abnormal conditions with computer aid.

Dougherty, E.M., 1983. Human reliability: Issues and practices in PRA. Transactions, American Nuclear Society, Vol. 45, 207-208.

Very brief discussion of techniques for human reliability analysis.

Embrey, D.E., 1981. The use of quantified expert judgement in the assessment of human reliability in nuclear power plant operations. Proceedings, Human Factors Society 25th Annual Meeting., 96-99.

Discusses the use of subjective judgement for reliability assessment.
Embrey, D.E., 1983. The use of performance shaping factors and quantified expert judgement in the evaluation of human reliability: an initial appraisal. NUREG 2986.

General discussion of HRA especially THERP. Quantifying error rates and discussion of the use of subjective judgement. Possible approaches for using expert judgement are discussed.

EPRI., 1977. Human factors review of nuclear power plant control room design. EPRI NP-309.

Chapter 19. "Operator errors. Discusses the "critical incident technique".

Farley, A.M., 1980. Issues in knowledge-based problem solving. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-10, No. 8.

Outlines a general model for knowledge-based problem solving including coordination of multiple plans, and time-related issues.

Finnegan, J.P., 1980. Workshop proceedings: The role of personnel error in fossil fuel power plant equipment reliability. EPRI AF-1470.

Attempts to discover the causes of personnel error and lists types of errors that have been observed.

Finnegan, J.P.. Rettig, T.W., and Rau, Jr., C.A., 1979. The role of personnel errors in power plant equipment reliability. EPRI AF-1041.

Interviews, questionnaires, analysis of trouble memos, and a survey of 38 fossil fuel power plant provided preliminary quantification of the effect of human error on availability.

Flathers II, G.W., Giffin, W.C., and Rockwell, T.H., 1982. A study of decision-making behaviour of aircraft pilots deviating from a planned flight. Aviation, Space, and Environmental Medicine, Vol. 53(10), 958-963.

Discusses analysis of behaviour under unexpected event conditions,. especially the 'worth' factor in the decision-making process. Behaviour:
1) detection 2) Diagnosis 3) decision 4) execution.

Gai, E.G., and Curry, R.E., 1976 A model of the human observer in failure detection tasks. IEEE Transactions on Systems,Man and Cybernetics, Vol. SMC-6, No. 2.

Statistical pattern recognition model for high frequency phenomena.

Greenstein. J.S. and Rouse, W.B., 1982. A model of human decision-making in multiple process monitoring situations. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, no. 2.

Event detection, monitoring and queueing theory based on division of attention - single channel information processor.

Hall, R.E., Fragola, J., and Wreathall, J., 1982. Post event human decision errors: Operator action tree / time reliability correlation. NUREG 3010.

The use of Operator Action Trees (OAT) for estimating human reliability. Concerned with errors due to diagnosis (thus different from THERP which is concerned with skill- and rule-based levels of behaviour). Defines and discusses cognitive behaviours, identification, decision processes, mental models. Time is a critical factor in success, 3 kinds of errors for diagnosing and correcting an abnormal event. Quantification and methodology for the accident environment.

Halpin, S.M., Johnson, E.M., and Thornberry, J.A., 1973. Cognitive reliability in manned systems. IEEE Tranactions, Reliability, Vol. R-22, No. 3, 165-170.

Discusses cognitive reliability and the types of error that might be included in this category.

Henneman, R.L., and Rouse, W.B., 1984. Measures of human problem solving performance in fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-14, no. 1.

Hollnagel, E., 1978. Diagnostic behaviour as conflict resolution. RISO-M-2110.

Detection, prediction, forecasting, selection of action. Model with 2 characteristics presents 4 situations the opoerator can potentially face.

Hollnagel, E., 1983. What we do not know about man-machine systems. Journal of Man-Machine Studies, 18, 135-143.

The human element in complex systems, performance shaping factors, strategies, and decision-making.

Hollnagel, E., and Woods, D.D., 1983. Cognitive systems engineering: New wine in new bottles. International Journal of Man-Machine Studies, 18, 583-600.

General discussion of cognitive engineering with discussion of error, human machine interaction.

Hunt, R.M., and Rouse, W.B., 1981. Problem-solving skills of maintenance trainees in diagnosing faults in simulated powerplants. Human Factors, 23(3), 317-328.

Hunt, R.M. and Rouse, W.B., 1984. A fuzzy rule-based model of human problem-solving. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-14, No. 1.

Huston, R.L., and Strauss, A.M., 1974. Human reliability in man-machine interactions. IEEE Proceedings, 1974 Annual Reliability Maintenance Conference, 329-334.

A mathematical model of the human response to machine demands. A sociomechanical approach concentrating on human conflict.

Jagacinski, R.J., and Miller, R.A., 1978. Describing the human operator's internal model of a dynamic system. Human Factors, Vol. 20, No. 4, 425-433.

Subjects' responses to a moving dot on a oscilloscope screen were measured and response patterns were used to infer their internal models. Describes learning behaviour and application to optimal control models.

Janis, I.L., and Mann, L., 1977. Emergency decision making: A theoretical analysis of responses to disaster warnings. Journal of Human Stress, 35-48.

Proposes a model which includes prediction and prognosis of consequences of actions or absence of actions. Describes 5 coping patterns, only one of which leads to appropriate decision-making.

Johannsen, G., and Rouse, W.B., 1983. Studies of planning behaviour of aircraft pilots in emergency and abnormal situations. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-13 No. 3.

Johnson, W.B. and Rouse, W.B., 1982. Analysis and classification of human error in troubleshooting live aircraft power plants. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-12, No. 3.

Error classification schemes. Looks at error following 3 different types of training evaluators identify errors on a checklist. Deals only with DIAGNOSIS. Lists types of errors that may occur when diagnosing.

Johnson-Laird, P.N., 1983. *Mental Models*, Cambridge: University Press.

On mental logic, drawing inferences, general topics in cognition, reasoning, problem solving.

Knaeuper, A.E., 1983. A model of human problem solving in dynamic environments. *Center of Man-Machine Systems Research*. Georgia Institute of Technology. Rept. No. 83-3.

Review of models of problem solving and development of a rule-based model (KARL) to simulate human performance.

Kozinsky, E.J., Gray, L.H., Beare, A.N., Barks, D.B. and Gomer, F.E., 1984. Safety-related operator actions: Methodology for developing criteria. *NUREG 3515*.

Objective was to develop a data base of quantitative measurements of operator actions under emergency conditions. Behaviour is in 4 phases:

1. Stimulus organization or observation.
2. Hypothesis generation, identification and interpretation.
3. Option selection or table definition.
4. Responses execution or output actions.

Discusses models of human performance, human reliability, human decision making Goal is to derive a simulation model of operator performance (OPPS).

Kragt, H., 1978. Human reliability engineering. *IEEE Transactions Reliability*, Vol. R-27, No. 3, 195-201.

General discussion of error. Taxonomy of behaviour in information processing terms and classification of techniques for HRE.

Kragt, H., and Landeweerd, J.A., 1974. Mental skills in process control. In: E. Edwards and F.P. Lees (Eds.). *The Human Operator in Process Control*. Taylor and Francis Ltd., London, 134-145.

Extracting the mental models of process controllers through interviews. Analyzing mental activities.

Kvalseth, T.O., 1977. A decision-theoretic model of the sampling behaviour of the human process monitor. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-7, No. 11, 810-813.

Application of decision theory to Senders' monitoring model.

Landeweerd, J.A., 1979. Internal representation of a process, fault diagnosis, and fault correction. *Ergonomics*, 22(12), 1343-1351.

Lees, F.P., 1973. Quantification of man-machine system reliability in process control. *IEEE Transactions Reliability*, Vol. R-22, No. 3, 124-131.

Discussion of reliability, emergency behaviour, and a taxonomy for reliability of process control tasks.

Lees, R.P. and Sayers, B., 1976. The behaviour of process operators under emergency conditions. In T.B. Sheridan and G. Johannsen (Eds.). *Monitoring Behaviour and Supervisory Control*. NATO Conference Series: III, Human Factors; v. 1) Plenum press, New York.

Reports experimental work on emergency behaviour, simulations of reactor faults.

Lund, L.O. and Sherwood, C., 1981. Non-obtrusive techniques for monitoring operator performance in nuclear power plant simulators. *Proceedings Human Factors Society 25th Annual Meeting*, 23-29.

Interviews, observation techniques of operators interacting with NPP simulators to assess cognitive functioning.

McLeod, P., 1976. Control strategies of novice and experienced controllers with a slow response system (a zero-energy nuclear reactor). In T.B. Sheridan and G. Johannsen (Eds.),. *Monitoring Behaviour and Supervisory Control*, NATO Conference Series: III, Human Factors; v. 1) Plenum Press, New York.

Investigation of strategies on a simulated reactor.

Meister, D., 1973. A critical review of human performance reliability predictive methods. *IEEE Transactions on Reliability*, Vol. R-22, No. 3, 116-123.

Reviews 22 methods of predicting operator reliability.

Meister, D., 1984. Human reliability. Human Factors Review, 13-53.

Definition and importance of human reliability. A historical review of earlier work plus methods, simulation models, and theoretical basis. Also a discussion of taxonomies.

Morris, M.N., 1983. Human problem solving in process control, Center for Man-Machine Systems Research. Georgia Institute of Technology, Rep. No. 83-2.

Literature review of process control. Simulations of fault detection tasks were used to investigate the variables affecting performance.

Nawrocki, L.H., Strub, M.H., and Cecil, R.M., 1973. Error categorization and analysis in man-computer communications systems. IEEE Transactions on Reliability, Vol R-22.

Discusses traditional approaches to human reliability and presents a method of classifying operator errors in human-computer systems.

Norman, D.A., 1980. Errors in human performance. UCSD Rep. No. 8004.

The nature of error and the role of knowledge. Discussion of TMI; summary of existing literature on error and its classification.

Norman, D.A., 1981. Categorization of action slips. Psychological Review, Vol. 88, No. 1.

General discussion of unintentional errors including a theory of action.

Norman, D.A., 1983. On human error: Misdiagnosis and failure to detect the misdiagnosis (Draft).

General discussion of misdiagnosis with emphasis on NPPs.

Norman., D.A., 1983(a). Design rules based on analyses of human error. Communications of the ACM, Vol. 26, No. 4., 254-258.

Applying knowledge about error to the design of systems; especially computer interface.

Norman, D.A., 1983(b). Position paper on human error. NATO Conference on Human Error, Bellagio, Italy.

Definition, taxonomy, theory and prediction of error.

Pederson, O.M. and Rasmussen, J. Mechanisms of human malfunction. Definition of categories: Revisions and Examples. RISO Electronics Dept. N-27-80.

Takes examples from nuclear power plants and attempts to classify each according to a malfunction mechanism.

Pew, E.W., Miller, D.C., and Feeker, C.E., 1981. Evaluation of proposed control room improvements through analysis of critical operator decisions. EPRI NP-1982.

Decision-making in abnormal events. Based on the Rasmussen model. Analysis of decision-making under emergency conditions.

Pew, E.W. and Baron, S., 1983. Perspectives on human performance modelling. Automatica, Vol. 19, No. 6, 663-675

Deals with performance modelling in general, i.e. what is a model? Why do it? Provides a categorization of models.

Potash, L.M., Stewart, M., Dietz, P.E., Levis, C.M. and Dougherty, E.M., Jr., 1981.

Experience in integrating the operator contributions in the PRA of actual operating plants. Proceedings of the International ANS/ens Topical Meeting on Probabilistic Risk Assessment, 1054-1063. Long Island Section: American Nuclear Society. (as cited in Norman, 1983.)

Price, H.E., Maisano, R.E., and Van Cott, H.P., 1982. The allocation of functions in man-machine systems: A perspective and a literature review. NUREG-2623.

Outlines human-machine roles and the differences between them. Reviews literature for role allocation in complex systems. Decisionmaking in emergencies model for NPPs.

Rasmussen, J., 1976. Outlines of a hybrid model of the process plant operator. In: T.B. Sheridan and G. Johannsen (Eds.), Monitoring Behaviour and Supervisory Control. NATO Conference Series III, Human Factors; v. 1), Plenum Press, New York, 371-383.

Discussion of methods applicable to real-life tasks (protocol analysis). Modelling performance.

Rasmussen, J., 1977. What can be learned from human error reports. In Duncan, Gruneberg, and Wallis (eds.), Changes in Working Life, John Wiley and Sons.

General discussion of error and levels of behaviour.

Rasmussen, J., 1979. On the structure of knowledge -- a morphology of mental models in a man-machine system context. *RISO-M-2192*.

Theoretical discussion of levels of a system that humans represent and the implications of these levels for dealing with a complex system.

Rasmussen, J., 1980. Notes on human error analysis and prediction. In Postolakes, G., Garriga, S., and Volta, G. (eds.) *Synthesis and Analysis Methods for Safety and Reliability Studies*. Plenum Press, New York..

More general discussion - distinguishes random errors from systematic and discusses reliability.

Rasmussen, J., 1981. Human errors. A taxonomy for describing human malfunction in industrial installations. *RISO-M-2304*.

Discussion of error and types of error, causes, performance shaping factors, error prediction.

Rasmussen, J., 1982. Human factors in high risk technology. Human reliability in risk analysis. In A.E. Green (ed.), *High Risk Safety Technology*, John Wiley and Sons Ltd.

Definition of accident, anatomy of an accident, risk vs. reliability analysis, categories of behaviour and error.

Rasmussen, J., 1983(c). Human errors in process control. Position paper for *NATO workshop on the Origin of Human Error*, Bellagio, Italy.

General discussion of error definition, cause and theory.

Rasmussen, J., 1983(b). Skills, rules, and knowledge; Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-13, No. 3.

Rasmussen, J., and Lind, M., 1981. Coping with complexity. *RISO-M-2293*.

What is complexity and how do operators deal with it? Signals/signs/symbols distinction.

Rasmussen, J., Pederson, O.M., Mancini, G., Carniho, A., Griffon, M., and Gagnolet, P., 1981. Classification system for reporting events involving human malfunctions. *RISO-M-2240*.

Presents a classification system that attempts to identify the mechanism of malfunction in order to quantify error rate meaningfully.

Rouse, W.B., 1978(a). A model of the human in a cognitive prediction task. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-8, No. 5, 473-477.

Theoretical model relying on STM and LTM.

Rouse, W.B., 1978(b). Human problem solving performance in a fault diagnosis task. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-8, No. 4, 258-271.

Reviews the literature on problem solving. Investigates the effects of problem size, forced pacing, computer aiding, and training on problem solving.

Rouse, W.B., 1978(b). A model of human decisionmaking in a fault diagnosis task. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-8, No. 5, 357-361.

A fuzzy set model of decision making during troubleshooting. A simulation study comparing the model and human performance.

Rouse, W.B., 1979. A model of human decisionmaking in fault diagnosis tasks that include feedback and redundancy. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-9,No. 4.

An extension of the fuzzy set model in Rouse (1978) with the added revision of including the roles of feedback and redundancy.

Rouse, W.B., 1983. Models of human problem solving: Detection, diagnosis and compensation for system failures. *Automatica*, Vol. 19, No. 6, 613-625.

A general review of models of detection, diagnosis, and compensation. Rouse outlines an overall model incorporating possible error mechanisms.

Rouse, W.B. and Rouse, S.H., 1982. Cognitive style as a correlate of human problem solving performance in fault diagnosis tasks. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-12, No. 5, Sept./Oct.

Rouse, W.B. and Morris, N.M., 1981. Studies of human problem solving in a process control task. *Proceedings - IEEE International Conference of Cybernetics and Society*, 479-487.

3 conflicting goals: stabilization, optimization, and diagnosis. Involves decision-making; uses PLANT, a simulation of a process control task to set up problems.

Rouse, W.B. and Rouse, S.H., 1979. Measures of complexity of fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 11, 720-727.

Literature review on complexity. Which are the best measures for predicting fault diagnosis performance?

Rouse, W.B., Rouse, S.H. and Pellegrino, S.J., 1980. A rule-based model of human problem solving performance in fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-10, No. 7.

Set of heuristics for low frequency events.

Sargent, T.O., 1981. A study of human behvaiuor in adverse stress. Presented at the American Nuclear Society Annual Meeting, Bal Harbour, Florida.

Discusses theoretical models of behaviour under stress, group behaviour, physiological responses. Large bibliography.

Seaver, D.A., Stillwell, W.G., 1983. Procedures for using expert judgement to estimate human error probabilities in nuclear power plant operations. NUREG-2743.

Senders, J.W., 1983. Human error and human reliability in process control. Chemical Process Control, II.

Discussion of the nature of error.

Sheridan, T.B., 1980. Errors in nuclear power plants. Technology Review, 22-33.

Nature of human error, and a discussion of human reliability methods.

Sheridan, T.B., 1983. Measuring, modelling, and augmenting reliability of man-machine systems. Automatica, Vol. 19, No. 6, 637-645.

General discussion of error and reliability and modelling reliability.

Shum, D.A. and Pfeiffer, P.E., 1983. Observer reliability and human inference. IEEE Transactions on Reliability, Vol, R-22, No. 3, 170-176.

Formal mathematical model for making inferences from a collection of unreliable sensors. Based on probability theory.

Speaker, D.M., Voska, K.J. and Luckas, Jr., W.J., 1983. Identification and analysis of human errors underlying electrical/electronic component related events reported by nuclear power plant licensees. NUREG-2987.

Identification and distribution of errors during normal maintenance, operation, and testing of NPP. No rates, but attempts to pinpoint causes.

Swain, A.D. and Guttmann, H.E., 1983. Handbook of human reliability analysis with emphasis on nuclear power plant applications. NUREG-1278.

Chapter 12 is an interim model for the quantification of human error probabilities for cognitive errors or errors which might occur following a plant upset.

Taylor, D.H., 1976. Accidents, risks and models of explanation. Human Factors, 18(4), 371-380.

Defining accidents (4 types). Types of explanations: causal and purposive defining risk and risk-taking.

Terano, T., Murayama, Y.M, and Alciyama, N., 1983. Human reliability and safety evaluation of man-machine systems. Automatica, Vol. 19, No. 6, 719-722.

Propose a method of quantifying human reliability for cognition and judgement. Quantifies percent errors per individual for well-defined experimental tasks.

The Man-Machine Interface and Human Reliability. 1981. An assessment and Projection. Conference Record of the IEEE Standards Workshop on Human Factors and Nuclear Safety at Myrtle Beach, S.C..

Special reference to pp. 48, 81, 201, 216. Review of many issues in human error and reliability.

Topmiller, D.A., Eckel, J.S. and Kozinsky, E.J., 1982. Human reliability data bank for nuclear power plant operations. Vol 1: A review of existing human reliability data banks. NUREG-2744

Not relevant to cognitive modelling or processes but points out the need for a structured data base for human error rates.

van de Graaf, R.C. and Wewerinke, P.H. Experimental and theoretical analysis of human monitoring and decision-making behaviour in failure detection tasks. National Aerospace Laboratory NLR, The Netherlands, NLR MP 81032 U.

A linear estimation and classical sequential decision theory model was used to investigate the effects of a number of variables on response time, false alarm rates, etc.

van Eekhout, J.M., and Rouse, W.B., 1981. Human errors in detection, diagnosis, and compensation for failures in the engine control room of a supertanker. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-11, No. 12, 813-816.

Experts were observed coping with failures on a simulator. Error classification.

van Gelder, K., 1980. Human error with a blending process simulator. IEEE Transactions on Reliability, Vol. R-29, No. 3.

Simulation of a chemical batch process to investigate the effect of training and experience, and task instruction on operator error.

Veldhuyzen, W. and Stassen, H.G., 1977. The internal model concept: An application to modelling human control of large ships. Human Factors, 19(4), 367-380.

Discusses the importance of including the internal models or knowledge sets of th operators when simulating human behaviour. Loists HOs purposes in building internal models. Includes an equation to predict future ship states and a decision making element. Model incorporates the updating of the model.

Ward, S.L. and Poturalski, R.J., 1982. Changes in maze-solving errors due to stress. Proceedings of Human Factors Society Annual Meeting, 479-482.

Induced stress by limiting the time available to solve the problem. Errors of commission tended to increase with stress indicating a deterioration in planning abilities.

West, B. and Clarke, J.A., 1974. Operator interaction with a computer controlled distillation column. In: E. Edwards and F.P. Lees (Eds.), The Human Operator in Process Control. Taylor and Francis Ltd., London., 206-221.

Some data on operator behaviour supervising a computer controlled pilot plant - includes response to breakdown.

Wewerinke, P.H., 1983. Model of the human observer and decision maker -- theory and validation. Automatica, Vol. 19, No. 6, 693-696.

A model which incorporates decision speed/accuracy and attention for observing a dynamic system.

Wickens, C.D., and Kessel, C., 1979. The effects of participatory mode and task workload on the detection of dynamic system failures. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 1, 224-34.

Detection performance and the variables affecting it (participation, concurrent task load).

Wickens, C.D. and Kessel, C., 1980. Processing resource demands of failure detection in dynamic systems. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 6(3), 564-577.

Used secondary task technique to measure the load imposed by manual control vs. monitoring. Found structure-specific resources involved.

Wohl, J.G., 1982. Maintainability prediction revisited: Diagnostic behaviour, system complexity, and repair time. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-12, No. 3.

Wreathall, J., 1983. Human reliability analysis: Fact or fiction? *Trans. Am. Nucl. Soc.*, Vol. 45, 209-210.

Brief discussion of methods for analysis of human reliability. Points out the need to consider cognitive aspects of human behaviour.

9.0 REFERENCES

Abo El-Ela, M.A., and Raouf, A., 1979. Manual backup operations: Some behavioural aspects of human reliability. Microelectronic Reliability., Vol. 19, 141-149.

Adams, J.A., 1982 Issues in human reliability. Human Factors, 24:1,. 1-10.

Altman, J.W., 1967. Classification of human error. In Symposium on reliability of human performance in work. W.B. Askren (Ed.). 1966 Annual Convention of the American Psychological Association. AMRL-TR-67-88.

Anyakora, S.N., and Lees, F.P., 1974 Detection of instrument malfunction by the process operator. In E. Edwards and F.P. Lees (Eds.), The Human Operator in Process Control. Taylor and Francis Ltd., London, 238-248.

Askren, W.B., and Regulinski, T.L., 1969. Quantifying human performance for reliability analysis of systems. Human Factors, 11(4), 393-396.

Bainbridge, L., 1974. Analysis of verbal protocols from a process control task. In Edwards, E., Lees, F. (Eds.), The Human Operator in process control. Taylor and Francis, London, 146-158.

Bainbridge, L., 1978(a). Forgotten alternatives in skill and workload. 169-185.

Bainbridge, L., 1978(b). The process controller. In Singleton, T., (Ed.), The study of real skills. Academic Press, N.Y.

Bainbridge, L., 1979. Verbal reports as evidence of the operator's knowledge. International Journal of Man-Machine Studies. 11, 411-436.

Bainbridge, L., 1981(a). Experimental studies and mathematical models of human problem solving performance in fault diagnosis tasks. In Rasmussen, J. and Rouse, W.B. (Eds.), Human detection and diagnosis of system failures. Plenum, New York, 199-216.

Bainbridge, L., 1981(b). Mathematical equations or processing routines? In Rasmussen, J. and Rouse, W.B. (Eds.), Human detection and diagnosis of system failures. Plenum Press, N.Y.

Bainbridge, L., Beishon, J., Hemming, J.H., and Splaine, M.A., 1974. A study of real-time human decision making using a plant simulator. In Edwards, E. and Lees, F. (eds.), The human operator in process control, Taylor and Francis, London. 90-104.

Baker, C., 1963. Further towards a theory of vigilance. In Buckner, D., and McGrath, J. (eds.), Vigilance: a symposium. McGraw-Hill, N.Y. 127-153.

Baron, S., Feehrer, C., Muraldhavan, R., Pew, R., Horwitz, P., 1982. An approach to modelling supervisory control of nuclear power plant. NUREG/CR-2988. Washington, D.C.

Baxa, E.G. and Nolte, L.W., 1972. Adaptive signal detection with finite memory. IEEE Transactions on Systems, Man and Cybernetics, SMC-2, 42-44.

Beare, A.N., Crowe, D.S., Kozinsky, E.J., Barks, D.B., Haas, P., 1982. Criteria for safety-related nuclear power plant operator actions: initial boiling water reactor (BWR) simulator exercises. NUREG/CR-2534, Washington, D.C.

Beare, A.N, Dorris, R.E., Bovell, C.R., Crowe, D.S., and Kozinsky, E.J., 1984. A simulator-based study of human errors in nuclear power plant control room tasks. NUREG/CR-3309.

Beare, A.N., Dorris, R.E., Kozinsky, E.J., Manning, J.J. and Haas, P., 1983. Criteria for safety-related nuclear power plant operator actions: initial simulator to field data calibration. NUREG/CR-3092, Washington, D.C.

Beattie, I.D and Iwasa-Madge, K.M., 1984. A Taxonomy for Human Reliability Analysis. Proceeding of the International Conference on Occupational Ergonomics, Toronto, Canada.

Bechtel, F.K. and Amos, D.E., 1975. Tradeoffs between detection probability and processing complexity. IEEE Transactions on Systems, Man and Cybernetics, SMC-5, 446-452.

Bell, B.J., and Swain, A.D., 1981. A procedure for conducting a Human reliability analysis for nuclear power plants. NUREG CR/2254.
Berliner, C. Angel, D., and Shearer, J.W., 1964. "Behaviours, Measures and Instruments for Performance Evaluation in Simulated Environments," Symposium and Workshop on the Quantification of Human Performance, Albuquerque, New Mexico.

Bisseret, A., 1971. Analysis of mental processes involved in air traffic control. Ergonomics, Vol. 14(5), 565-570.

Braune, R., and Forshay, W.R., 1983. Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. Instructional Science 12, 121-145.

Braune, R.J., and Trollip, S.R., 1982. Toward an internal model in pilot training. Aviation, Space, and Environmental Medicine, Vol. 53(10), pp. 996-999.

Bruner, J.S., 1973. Beyond the Information Given. Norton, New York.

Chu, Y.Y., Steeb, R., and Freedy, A., 1980. Analysis and modelling of information handling tasks in supervisory control of advanced aircraft. Perceptronics Inc. Tech. Rep. PATR-1080-80-6.

Coeterier, J.F., 1971. Individual strategies in ATC freedom and choice. Ergonomics, 14, 579-584.

Comer, M.K., Kozinsky, E.J., Eckel, J.S., and Miller, D.P., 1982 Human reliability data bank for nuclear power plant operations. Vol 2: NUREG/CR-2744, 2 of 2.

Crowe, D.S., Beare, A.N., Kozinsky, E.J., and Haas, P.M., 1982. Criteria for Safety Related Nuclear Power Plant Operator Actions: Pressurized Water Reactor (PWR) Simulator Exercises., NUREG/CR-3123.

Curry, R.E. and Nagel, D.C., 1976. Decision behaviour with changing signal strengths. Journal of Mathematical Psychology, 14, 1-24.

Dale, H., 1968. Weighing evidence: an attempt to assess the efficiency of the human operator. Ergonomics, 11, 215-230.

De Keyser, V., 1980. Le comportement humain dans la surveillance et la conduite des systemes automatiques. Proceedings of IFAP, Rome, 1980.

de Jong, J.J. and Koster, E.P., 1974. The human operator in the computer-controlled refinery. In: E. Edwards and F.P. Lees (Eds.). The Human Operator in Process Control Taylor and Francis Lts., London, 196-205.

Dougherty, E.M., 1983. Human reliability: Issues and practices in PRA. Transactions, American Nuclear Society, Vol. 45, 207-208.

Edwards, E. and Lees, F., 1974. The human operator in process control. Taylor and Francis, London.

Embrey, D.E., 1981. The use of quantified expert judgement in the assessment of human reliability in nuclear power plant operations. Proceedings, Human Factors Society 25th Annual Meeting. 96-99.

Embrey, D.E., 1983. The use of performance shaping factors and quantified expert judgement in the evaluation of human reliability: an initial appraisal. NUREG/CR-2986.

Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B., and Rea, K., 1984. SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgement. Vols. I and II. NUREG/CR-3518, Washington, D.C.

EPRI, 1977. Human factors review of nuclear power plant control room design. EPRI NP-309.

Eriksen, E., Simon, H., 1984. Protocol Analysis. MIT Press. Cambridge.

Farley, A.M., 1980. Issues in knowledge-based problem solving. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-10, No. 8.

Finnegan, J.P., 1980. Workshop proceedings: The role of personnel error in fossil fuel power plant equipment reliability. EPRI AF-1470.

Finnegan, J.P.. Rettig, T.W., and Rau, Jr., C.A., 1979. The role of personnel errors in power plant equipment reliability. EPRI AF-1041.

Flathers II, G.W., Giffin, W.C., and Rockwell, T.H., 1982. A study of decision-making behaviour of aircraft pilots deviating from a planned flight. Aviation, Space, and Environmental Medicine, Vol. 53(10), 958-963.

Gai, E.G. and Curry, R.E., 1976. A model of the human observer n failure detection tasks. IEEE Transactions on Systems, Man and Cybernetics, SMC-6, 85-91.

Gaines, B.R., 1976. On the complexity of causal models. IEEE Transactions on Systems, Man and Cybernetics, SMC-6, 56-59.

Gentner, D. and Stevens, A.L., 1983. Mental Models. Lawrence Erlbaum, New Jersey.

Gonzales, R.C. and Hovington, L.C., 1977. Machine recognition of abnormal behaviour in nuclear reactors. IEEE Transactions on Systems, Man and Cybernetics, SMC-7, 717-728.

Green, T.R.G., 1980. Human interaction with computers. Academic Press, London.

Greenstein. J.S. and Rouse, W.B., 1982. A model of human decision-making in multiple process monitoring situations. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, no. 2.

Hall, R.E., Fragola, J.R. and Wreathall, J., 1982. Post event human decision errors: Operator Action Tree/Time Reliability correlation. NUREG/CR-3010, Washington, D.C.

Halpin, S.M., Johnson, E.M., and Thornberry, J.A., 1973. Cognitive reliability in manned systems. IEEE Tranactions, Reliability, Vol. R-22, No. 3, 165-170.

Hannaman, G.W. and Spergin, A.J., 1984. Systematic Human Action Reliability Procedure (SHARP). EPRI-NP-3583, Palo Alto, Ca.

Hanneman, G.W., Spergin, A.J., and Lukic, 1984.

Henneman, R.L. 1981. Measures of human performance in fault tasks, M.S.I.E Thesis University of Illinois a Urbana-Champaign.

Henneman, R.L., and Rouse, W.B., 1984. Measures of human problem solving performance in fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-14, no. 1.

Hollnagel, E., 1978. Diagnostic behaviour as conflict resolution. RISO-M-2110.

Hollnagel, E., 1983. What we do not know about man-machine systems. Journal of Man-Machine Studies, 18, 135-143.

Hollnagel, E., and Woods, D.D., 1983. Cognitive systems engineering: New wine in new bottles. International Journal of Man-Machine Studies, 18, 583-600.

Hopf-Weichel, R., Lucaccini, L., Saleh, J., and Freedy, A. Aircraft emergency decisions: cognitive and situational variables. Perceptronics Inc. Tech Rept. PATR-1065-79-7.

Hunt, R.M., and Rouse, W.B., 1981. Problem-solving skills of maintenance trainees in diagnosing faults in simulated powerplants. Human Factors, 23(3), 317-328.

Hunt, R.M. and Rouse, W.B., 1984. A fuzzy rule-based model of human problem-solving. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-14, No. 1.

Huston, R.L., and Strauss, A.M., 1974. Human reliability in man-machine interactions. IEEE Proceedings, 1974 Annual Reliability Maintenance Conference, 329-334.

Iosif, G., 1968. La strategie dans la surveillance des tableaux de commande. I. Quelques facteurs determinants de caractere objectif. Revue Roumanien des Science Sociales -- Psychologique, 12, 147-161.

Iosif, G., 1969. La strategie dans la surveillance des tableaux de commande. II. Quelques facteurs determinants de caractere subjectif. Revue Roumanien de Science Sociale -- Psychologique, 13, 29-41. (a)

Iosif, G., 1969. Influence de la correlation fonctionelle sur parametres technologiques., Revue Roumenien de Science Sociale-Psychologique, 13, 105-110. (b)

Jagacinski, R.J., and Miller, R.A., 1978. Describing the human operator's internal model of a dynamic system. Human Factors, Vol. 20, No. 4, 425-433.

Janis, I.L., and Mann, L., 1977. Emergency decision making: A theoretical analysis of responses to disaster warnings. Journal of Human Stress, 35-48.

Johannsen, G., and Rouse, W.B., 1983. Studies of planning behaviour of aircraft pilots in emergency and abnormal situations. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-13 No. 3.

Johnson, W.B. and Rouse, W.B., 1982. Analysis and classification of human error in troubleshooting live aircraft power plants. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, No. 3.

Johnson-Laird, P., 1983. Mental Models. Harvard University Press, Cambridge, Ma.

Johnson-Laird , P.N., and Wason, P.C., 1970. A theoretical analysis of insight into a reasoning task. Cognitive Psychology, 1, 134-148.

Kay, H., 1950. Studies of memory. Unpublished Ph.D. thesis, University of Oxford, 1950.

Kemeny, J.G., 1979. Report of the President's commission on the accident at Three Mile Island. Pergamon Press, N.Y.

Kollarits, J., 1937. Bobachtungen uber, dysproxian (Fehlhandlungen). Archiv Fur Die Gesamte Psychologie, 99, 305-399.

Kozinsky, E.J., Gray, L.H., Beare, A.N., Barks, D.B. and Gomer, F.E., 1984. Safety-related operator actions: Methodology for developing criteria. NUREG/CR-3515.

Knaeuper, A.E., 1983. A model of human problem solving in dynamic environments. Center of Man-Machine Systems Research. Rept. No. 83-3. Georgia Institute of Technology.

Kragt, H., 1978. Human reliability engineering. IEEE Transactions Reliability, Vol. R-27, No. 3, 195-201.

Kragt, H., and Landeweerd, J.A., 1974. Mental skills in process control. In: E Edwards and F.P. Lees (Eds.). The Human Operator in Process Control. Taylor and Francis Ltd., London, 134-145.

Kvalseth, T., 1978. Human and Bayesian information processing during probabilistic inference tasks. IEEE Transactions on Systems, Man and Cybernetics, SMC-8, 224-229.

Kvalseth, T.O., 1977. A decision-theoretic model of the sampling behaviour of the human process monitor. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-7, No. 11, Nov. 810-813.

Landeweerd, J.A., 1979. Internal representation of a process, fault diagnosis, and fault correction. Ergonomics, 22(12), 1343-1351.

Lane, N.M., Strieb, Glenn, F.A., and Wherry, R.J., 1980. The human operator simulator: an overview. In Proceedings of Conference on Manned Systems Design, Freiburg, pp. 1-39.

Lees, F.P., 1973. Quantification of man-machine system reliability in process control. IEEE Transactions Reliability, Vol. R-22, No. 3, 124-131.

Lees, F.P. and Sayers, B., 1976. The behaviour of process operators under emergency conditions. In T.B. Sheridan and G. Johannsen (Eds.). Monitoring Behaviour and Supervisory Control. NATO Conference Series: III, Human Factors; v. 1) Plenum press, New York.

Levine, J. and Samet, M., 1973. Information seeking with multiple sources of conflicting and unreliable information. Human Factors Vol. 15, 407-419.

Levine, J, Samet, M and Brahlek, R., 1975. Information seeking with limitations on available information and resources.Human Factors, Vol. 17, 502-513.

Loftus, E.F., 1979. Eyewitness testimony. Harvard University Press, Cambridge, Mass.

Lund, L.O. and Sherwood, C., 1981. Non-obtrusive techniques for monitoring operator performance in nuclear power plant simulators. Proceedings Human Factors Society 25th Annual Meeting, 23-29.

Mackworth, J., 1959. Paced memorizing in a continuous task. Journal of experimental psychology, 58, 206-211.

Mason, J.F., 1979. The technical blow-by-blow. Spectrum, 1979, 11, 33-57.

McLeod, P., 1976. Control strategies of novice and experienced controllers with a slow response system (a zero-energy nuclear reactor). In T.B. Sheridan and G. Johannsen (Eds.),. Monitoring Behaviour and Supervisory Control, NATO Conference Series: III, Human Factors; v. 1) Plenum Press, New York.

Meister, D., 1983. A critical review of human performance reliability predictive methods. IEEE Transactions on Reliability, Vol. R-22, No. 3, 116-123.

Meister, D., 1984. Human reliability. Human Factors Review, 13-53.

Morris, M.N., 1983. Human problem solving in process control, Center for Man-Machine Systems Research. Rep. No. 83-2. Georgia Institute of Technology.

Moray, N., 1981. The role of attention in the detection of errors and the diagnosis of failures in man-machine systems. In Rasmussen, J. and Rouse, W.B. (eds.), Human detection and diagnosis of system failures. Plenum Press, N.Y., 185-198.

Moray, N., 1985. Monitoring behaviour and supervisory control. In Handbook of Human Perception and Performance (Eds. Boff, K., Kauffman, L. and Beattie, J.) Wiley, N.Y. (In press)

Mullis, C.T. and Roberts, R.A., 1968. Memory limitation and multistage decision processes. IEEE Transactions on Systems, Man and Cybernetics, SMC-4, 307-316.

Nawrocki, L.H., Strub, M.H., and Cecil, R.M. Error categorization and analysis in man-computer communications systems. IEEE Transactions on Reliability, Vol R-22, no. 3, 135-

Norman, D.A., 1980. Errors in human performance. UCSD Rep. No. 8004.

Norman, D.A., 1981. Categorization of action slips. Psychology Review, Vol. 88, No. 1.

Norman, D.A., 1983(a). Design rules based on analyses of human error. Communications of the ACM, Vol. 26, No. 4, April 254-258.

Norman, D.A., 1983(b). Position paper on human error. NATO Conference on Human Error, Bellagio, Italy, Sept.

Norman, D.A., 1983. On human error: Misdiagnosis and failure to detect the misdiagnosis (Draft).

NUREG/CR-3114, 1982. Proceedings of Workshop on Cognitive Modelling of Nuclear Plant Control Room Operators, August 15-18.

Pederson, O.M. and Rasmussen, J. Mechanisms of human malfunction. Definition of categories: Revisions and Examples. RISO Electronics Dept. N-27-80.

Peterson, C. and Beach, L., 1964. Man as an intuitive statistician. Psychological Bulletin, 68, 29-46.

Pew, R.W. and Baron, S., 1983. Perspectives on human performance modelling. Automatica, Vol. 19, No. 6, 663-675.

Phatak, A.V. and Bekey, G.A., 1969. Decision processes in the adaptive behaviour of human controllers. IEEE Transactions on System Science and Cybernetics, SSC-5, 339-352.

Phillips, L.D., Humphreys, P. and Embrey, D.E., 1983. A socio-technical approach to assessing human reliability. TR-83-4, ORNL.

Potash, L.M., Stewart, M., Dietz, P.E., Levis, C.M. and Dougherty, E.M., Jr., 1981. Experience in integrating the operator contributions in the PRA of actual operating plants. Proceedings of the International ANS/ens Topical Meeting on Probabilistic Risk Assessment, 1054-1063. Long Island Section: American Nuclear Society. (as cited in Norman, 1983.)

Price, H.E., Maisano, R.E., and Van Cott, H.P., 1982. The allocation of functions in man-machine systems: A perspective and a literature review. NUREG-2623,

Rasmussen, J., 1976. Outlines of a hybrid model of the process plant operator. In: T.B. Sheridan and G. Johannsen (Eds.), Monitoring Behaviour and Supervisory Control. NATO Conference Series III, Human Factors; v. 1), Plenum Press, New York, 371-383.

Rasmussen, J., 1977. What can be learned from human error reports. In Duncan, Gruneberg, and Wallis (eds.), Changes in Working Life, John Wiley and Sons.

Rasmussen, J., 1979. On the structure of knowledge -- a morphology of mental models in a man-machine system context. RISO-M-2192.

Rasmussen, J., 1980. Notes on human error analysis and prediction in Postolakes, G., Garriga, S., and Volta, G. (Eds.) Synthesis and Analysis Methods for Safety and Reliability Studies. Plenum. N.Y.

Rasmussen, J., 1981. Human errors. A taxonomy for describing human malfunction in industrial installations. RISO-M-2304.

Rasmussen, J., 1982. Human factors in high risk technology. Human reliability in risk analysis. In A.E. Green (ed.), High Risk Safety Technology, John Wiley and Sons Ltd.

Rasmussen, J., 1983(a). Design for Error Tolerances. Invited Summary for American Nuclear Society, Winter Meeting.

Rasmussen, J., 1983(b). Skills, rules and knowledge; signals, signs and symbols and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, SMC-13, 257-266.

Rasmussen, J., 1983(c). Human errors in process control. Position paper for NATO workshop on the Origin of Human Error, Bellagio, Italy.

Rasmussen, J., and Lind, M., 1981. Coping with complexity. RISO-M-2293.

Rasmussen, J., Pederson, O.M., Mancini, G., Carniho, A., Griffon, M., and Gagnolet, P., 1981. Classification system for reporting events involving human malfunctions. RISO-M-2240.

Rasmussen, J. and Rouse, W.B., 1981. Human detection and diagnosis of system failures. Plenum, New York.

Reason, J.T., 1977. Skill and Error in Everyday Life. In: Adult Learning, Editior M. Howe, London, Wiley.

Rolt, L.T.C., 1978. Red for Danger. Pan Books. London.

Rouse, W.B., 1973(a). Models of man as a suboptimal controller. 9th NASA Annual conference on Manual Control. (NTIS No. N75-19156/2.) 413-415.

Rouse, W.B., 1973(b). A model of the human in a cognitive prediction task. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-3, No. 5, 473-477.

Rouse, W.B., 1977. A theory of human decision making in stochastic estimation tasks. IEEE Transactions on Systems, Man and Cybernetics, SMC-7, 274-284.

Rouse, W.B., 1978(a)., Human problem solving performance in a fault diagnosis task. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-8, No. 4, 258-271.

Rouse, W.B., 1978(b)., A model of human decision-making in a fault diagnosis task. IEEE Transactions on Systems, Man, and Cybernetics, SMC-8, 357-361.

Rouse, W.B., 1979. A model of human decisionmaking in fault diagnosis tasks that include feedback and redundancy. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9,No. 4.

Rouse, W.B., 1981. Experimental studies and mathematical models of human problem solving performance in fault diagnosis tasks. In Rasmussen, J. and Rouse, W.B. (Eds.), Human detection and diagnosis of system failures. Plenum, New York, 199-216.

Rouse, W.B., 1983. Models of human problem solving: Detection, diagnosis and compensation for system failures. Automatica, Vol. 19, No. 6, 613-625.

Rouse, W.B. and Morris, N.M., 1981. Studies of human problem solving in a process control task. IEEE Intl. Conf. on Cyb. Soc., 479-487.

Rouse, W.B. and Rouse, S.H., 1979. Measures of complexity of fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 11, nov. 720-727.

Rouse, W.B. and Rouse, S.H., 1982. Cognitive style as a correlate of human problem solving performance in fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, No. 5.

Rouse, W.B., Rouse, S.H. and Pellegrino, S.J., 1980. A rule-based model of human problem solving performance in fault diagnosis tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-10, No. 7.

Rubenstein, E., 1979. The accident that shouldn't have happened. Spectrum, 16, 33-57.

Sage, A.P., 1981. Behavioural and organizational considerations in the design of information systems and processes for planning and decision support. IEEE Transactions on Systems, Man and Cybernetics, SMC-11, 640-678.

Schrenk, I.P., 1969. Aiding the decision maker: a decision process model. Ergonomics, 12, 543-557.

Schurman, D.L. and Banks, W.W., 1985. Review of Human Error Assessment Methods and Models. SCE-NTPD-84-003, EG & G Idaho Falls.

Sargent, T.O., 1981. A study of human behvaiour in adverse stress. Presented at the American Nuclear Society Annual Meeting, Bal Harbour, Florida.

Seaver, D.A., and Stillwell, W.G., 1983. Procedures for using expert judgement to estimate human error probabilities in nuclear power plant operations. NUREG-2743.

Sellen, A.J., 1983. The Frequency and Distribution of Errors Produced Under Time Stress. B.A. dissertation, University of Toronto, Department of Psychology.

Senders, J.W., 1983. Human error and human reliability in process control. Chemical Process Control, II.

Shepard, R.N., 1967. Recognition memory for words, sentences, and pictures. Journal of verbal learning and verbal behaviour, 6, 156-163.

Sheridan, T.B., 1980. Errors in nuclear power plants. Technology Review, 22-33.

Sheridan, T.B., 1982. Cognitive Models and Computer Aids for Nuclear Plant Control Room Operators. In Proceedings of Workshop on Cognitive Modelling of Nuclear Plant Control Room Operators, August 15-18.

Sheridan, T.B., 1983. Measuring, modelling, and augmenting reliability of man-machine systems. Automatica, Vol. 19, No. 6, 637-645.

Sheridan, T.B. and Ferrell, R., 1974. Man-machine systems. M.I.T. Press. Cambridge, Mass.

Sheridan, T.B., and Johannsen, G., 1976. Monitoring behaviour and supervisory control. Plenum. N.Y.

Shiff, B., 1983. An experimental study of the human-computer interface in process control. Unpublished M.A.Sc. thesis, Department of Industrial Engineering, University of Toronto.

Shum, D.A. and Pfeiffer, P.E., 1973. Observer reliability and human inference. IEEE Transactions on Reliability, Vol, R-22, No. 3, Aug. 170-176.

Siegal, A.I., Bartter, W.D., Knee. H.E. and Haas, P.M., 1983. Front-end analysis for the nuclear power plant maintenance personnel reliability model. NUREG/CR-2669, Washington, D.C.

Siegal, A.I. and Wolf, J.J., 1967. Man-machine simulation models: performance and psychological interaction. Wiley, New York.

Soede, M. and Coeterier, J., 1971. Time analysis of tasks of approach controllers in ATC. Ergonomics, 14, 511-517.

Speaker, D.M., Voska, K.J. and Luckas, Jr., W.J., 1983. Identification and analysis of human errors underlying electrical/electronic component related events reported by nuclear power plant licensees. NUREG/CR-2987.

Sperandio, J.C., 1971. Variation of operators' strategies and regulating effects of workload. Ergonomics, 14, 511-517.

Sperandio, J.C., 1978. The regulation of working methods as a function of workload among air traffic controllers. Ergonomics, 21, 195-200.

Swain, A.D. and Guttmann, H.E., 1982. Handbook of human reliability analysis with emphasis on nuclear power plant applications. NUREG/CR-1278, Washington, D.C.

Taylor, D.H., 1976. Accidents, risks and models of explanation. Human Factors, 18(4), 371-380.

Taylor, F.J., 1975. Finite fading memory filtering. IEEE Transactions on Systems, Man and Cybernetics, SMC-5, 134-137.

Terano, T., Murayama, Y.M, and Alciyama, N., 1983. Human reliability and safety evaluation of man-machine systems. Automatica, Vol. 19, No. 6, 719-722.

Topmiller, D.A., Eckel, J.S. and Kozinsky, E.J., 1982. Human reliability data bank for nuclear power plant operations. Vol 1: A review of existing human reliability data banks. NUREG/CR-2744.

Tversky, A. and Kahneman, D., 1971. The belief in the "law of small numbers". Psychological Bulletin, 76, 105-110.

van de Graaf, R.C. and Wewerinke, P.H., 1981. Experimental and theoretical analysis of human monitoring and decision-making behaviour in failure detection tasks. National Aerospace Laboratory NLR, The Netherlands, NLR MP 81032 U.

van Eekhout, J.M., and Rouse, W.B., 1981. Human errors in detection, diagnosis, and compensation for failures in the engine control room of a supertanker. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-11, No. 12, 813-816.

van Gelder, K., 1980. Human error with a blending process simulator. IEEE Transactions on Reliability, Vol. R-29, No. 3.

Vaughan, W.S. and Mavor, A.S., 1972. Behavioural characteristics of men in the performance of some decision-making task components. Ergonomics, 15, 267-277.

Veldhuyzen, W. and Stassen, H.G., 1977. The internal model concept: An application to modelling human control of large ships. Human Factors, 19(4), 367-380.

Ward, S.L. and Poturalski, R.J., 1982. Changes in maze-solving errors due to stress. Proceedings of Human Factors Society Annual Meeting, 479-482.

West, B. and Clarke, J.A., 1974. Operator interaction with a computer controlled distillation column. In: E. Edwards and F.P. Lees (Eds.), The Human Operator in Process Control. Taylor and Francis Ltd., London, 206-221.

Wason, P.C., 1968. Reasoning about a rule. Quarterly Journal of Experimental Psychology, 20, 273-281.

Wewerinke, P., 1981. A model of the human decision maker observing a dynamic system. Technical Report. NCR TR 81062 National Lucht - en Ruintevaart laboratorium, Netherlands.

Wewerinke, P.H., 1983. Model of the human observer and decision maker -- theory and validation. Automatica, Vol. 19, No. 6, 693-696

Wickens, C.D., and Kessel, C., 1979. The effects of participatory mode and task workload on the detection of dynamic system failures. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 1, 24-34.

Wickens, C.D. and Kessel, C., 1980. Processing resource demands of failure detection in dynamic systems. Journal of Experimental Psychology: Human Perception and Performance, Vol. 6(3), 564-577.

Wohl, J.G., 1982. Maintainability prediction revisited: Diagnostic behaviour, system complexity, and repair time. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, No. 3.

Wreathall, J., 1983. Human reliability analysis: Fact or fiction? Trans. Am. Nucl. Soc., Vol. 45, 209-210.

Yates, F., 1966. The art of memory. Routledge and Kegan Paul, London.

Zeitlin, L.R. and Finkleman, J.M., 1975. Research note: subsidiary task techniques of digit generation and digit recall as indirect measures of operator loading. Human Factors, 1975, 17, 218-220.

10.0 ABBREVIATIONS AND ACRONYMS USED IN TEXT

ATC	Air Traffic Control
BWR	Boiling Water Reactor
CANDU	Canadian Deuterium Uranium
CRT	Cathode Ray Tube
EMBED	Error Model Based on Empirical Data
EOP	Emergency Operating Procedure
EXHUME	Exhaustive Model of Human Error
FAULT	Context-specific training model. (Evaluated by Johnson & Rouse, 1982.)
HOS	Human Operator Simulator
HRA	Human Reliability Analysis
IEEE	Institute of Electrical and Electronic Engineers
LER	Licencee Event Report
LOCA	Loss Of Coolant Accident
MAPPS	Maintenance Personnel Performance Simulation
NPP	Nuclear Power Plant
OATS	Operator Action Trees
PORV	Pressure Operated Relief Valve
PRA	Probabilistic Risk Assessment
PROCRU	Procedure Oriented Crew Model
PSF	Performance Shaping Factors
PWR	Pressurized Water Reactor
S/N	Signal to Noise ratio
SAINT	Systems Analysis of Integrated Networks of Tasks
SDT	Signal Detection Theory
SHARP	Systematic Human Reliability Procedure

SER	Significant Event Report
SLI	Success Likelihood Index
SLIM-MAUD	Success Likelihood Index Methodology - Multi Attribute Utility Decomposition.
STAHR	Socio-Technical Assessment of Human Reliability
TASK	Context-free training method. (Evaluated by Johnson & Rouse, 1982.)
THERP	Technique For Human Error Rate Prediction
VDT	Video Display Terminal

11.0 Acknowledgements

The authors gratefully acknowledge the assistance E. Lindsay Gorrell, the AECB contract monitor.